N.YA. VILENKIN G.E. SHILOV
V.A. USPENSKII IU.I. LIUBICH

QUELQUES APPLICATIONS DES MATHÉMATIQUES

N. Vilenkine, G. Chilov, V. Ouspenski,
J. Lioubitch, L. Chor

QUELQUES APPLICATIONS DES MATHÉMATIQUES

MÉTHODE DES APPROXIMATIONS SUCCESSIVES

GAMME SIMPLE (STRUCTURE DE L'ÉCHELLE MUSICALE)

QUELQUES APPLICATIONS DE LA MÉCANIQUE AUX MATHÉMATIQUES

MÉTHODE CINÉMATIQUE DANS LES PROBLÈMES DE GÉOMÉTRIE

Editions Mir · Moscou

PRÉFACE

Le présent ouvrage est constitué de quatre livres d'auteurs différents, empruntés à la série « Cours de mathématiques », éditée à l'intention des lycéens. L'ordre de lecture importe peu. Si ces quatre « nouvelles » semblent sans rapport, elles n'en possèdent pas moins quelques affinités, dont la plus évidente est le refus délibéré des auteurs d'aborder les sujets traités dans un esprit « puriste », c'est-à-dire celui du mathématicien qui se refuse à constater les liens multiples et féconds existant entre les sciences mathématiques et celles de la nature.

Le premier livre: « Méthode des approximations successives » de N. Vilenkine, développe d'importantes questions des mathématiques « numériques », qui sont indispensables à tout chercheur et ingénieur, et qui, par la manière dont elles sont abordées, se démarquent nettement des mathématiques « démonstratives » des Grecs et de leurs disciples contemporains, bien plus attachés à la finesse des constructions de l'esprit qu'à la valeur pratique des résultats obtenus [1]).

Le deuxième livre: « Gamme simple » (structure de l'échelle musicale) est de G. Chilov. L'auteur étudie un problème qui, s'il remonte à l'Antiquité, n'a en fait été résolu qu'il y a trois siècles à peine, en l'occurrence: comment choisir les sons musicaux qui donneraient le plus de combinaisons harmonieuses. De par son énoncé ce problème est bien plus appliqué que « théorique » ne serait-ce que

---

[1]) Cette manière de concevoir des mathématiciens de l'Antiquité est rapportée dans une anecdote célèbre: Euclide aurait dit à son esclave à l'adresse d'un nouveau disciple, soucieux du profit qu'il pourrait tirer des théories étudiées: « Donne-lui de l'argent, car il veut tirer profit de l'étude des mathématiques ». (Remarquons qu'Euclide est l'auteur de travaux en optique, théorie de la réflexion, musique, astronomie, et il paraît peu probable qu'il ait manifesté un tel mépris envers ce désir de tirer profit des mathématiques.)

parce qu'il n'est pas formalisé et réduit à n'admettre qu'une seule solution (ainsi nous n'avons pas par exemple défini la notion exacte de « harmonieux »). Mais précisément ceci est très fréquent dans l'étude des problèmes appliqués pour lesquels le processus de formalisation (qui est inexistant dans les mathématiques « pures ») constitue souvent l'une des plus importantes étapes de la résolution.

Les troisième et quatrième livres, « Quelques applications de la mécanique aux mathématiques » d'Ouspenski et « Méthode cinématique dans les problèmes de géométrie » de Lioubitch et Chor, procèdent d'une démarche inverse dans l'étude des rapports traditionnels des mathématiques et de la mécanique en ce sens que leurs auteurs se penchent non pas sur l'application des mathématiques à la mécanique (telle qu'on la conçoit au niveau du secondaire), mais, au contraire, sur une interprétation mécanique des problèmes de mathématiques, en faisant intervenir des raisonnements propres à la physique. S'il est vrai que ces raisonnements « exogènes » sont rares, ils n'ont pas moins contribué dans un large domaine à élargir l'horizon de nos connaissances mathématiques : Bernard Riemann (1826-1866), célèbre mathématicien allemand, en tira le meilleur effet dans ses travaux. Pourtant, les raisonnements de la physique sont riches d'enseignements sur le plan des méthodes employées en ce sens qu'ils illustrent parfaitement les liens existant entre les mathématiques et les sciences de la nature, liens qui sont à l'origine d'une foule de notions et de théories mathématiques. Vus sous cet angle, ces deux derniers livres présentent un intérêt certain, même si les sujets traités semblent dépassés.

N. Vilenkine

Méthode des approximations successives

§ 1 INTRODUCTION

Les programmes scolaires réservent une grande place à la résolution des équations et systèmes d'équations: d'abord aux équations du premier degré et aux systèmes d'équations du premier degré, puis aux équations du second degré, bicarrées et irrationnelles, et, enfin, aux équations exponentielles, logarithmiques et trigonométriques.

Cette étude détaillée s'explique par l'importance des équations dans les applications mathématiques. Et pour cause: quel que soit le domaine considéré, la clef du problème est souvent donnée par la résolution d'une équation ou d'un système d'équations.

Au lycée on recourt souvent aux équations pour résoudre des problèmes de physique. Prenons un exemple:

On lâche une pierre dans un puits. On demande de trouver la profondeur de ce puits si le son provoqué par la chute met T secondes à nous parvenir.

Si nous désignons par x la profondeur du puits, nous obtenons l'équation suivante

$$\sqrt{\frac{2x}{g}}+\frac{x}{v}=T,$$

où v représente la vitesse de propagation du son dans l'air, $\sqrt{\frac{2x}{g}}$ la durée de la chute et $\frac{x}{v}$ le temps mis par le son pour nous parvenir. C'est une équation irrationnelle. Si l'on pose $\sqrt{x}=y$, elle se ramène à l'équation du second degré

$$\frac{y^2}{v}+\sqrt{\frac{2}{g}}\,y-T=0,$$

que l'on résout par une formule connue.

Les équations servent également à résoudre des problèmes de géométrie. Exemple: le problème qui consiste à trouver un point C partageant un segment AB de longueur l

dans un rapport k tel que $\frac{AB}{AC}=\frac{AC}{CB}=k$ se ramène à la résolution de l'équation

$$x^2 + lx - l^2 = 0,$$

où x désigne la longueur du segment AC.

La trisection d'un angle α conduit à une équation d'une forme plus compliquée:

$$4x^3 - 3x - \cos\alpha = 0,$$

où $x = \cos\frac{\alpha}{3}$. Cette équation dite *équation cubique* ne figure pas au programme du secondaire. La formule qui donne sa solution est établie dans les cours d'algèbre supérieure (cf. plus bas formule (3)).

La physique nous confronte souvent à des problèmes qui se ramènent à des équations d'une forme plus complexe, qui ne possèdent de formule de résolution toute prête ni dans les programmes du secondaire ni dans ceux du supérieur. Prenons par exemple une tige métallique fixée à demeure par ses extrémités. Si l'on déplace cette tige de sa position d'équilibre, elle se mettra à vibrer transversalement. En physique mathématique on démontre que la fréquence de ces vibrations obéit à l'équation

$$e^x + e^{-x} = \frac{2}{\cos x}, \qquad (1)$$

où $e = 2{,}71828...$

Dans le secondaire, il n'existe aucune formule donnant la solution de cette équation. Il ne faut pas croire que c'est une rançon de l'étendue du programme de mathématiques. Il n'existe pas de formule, au sens qu'on lui accorde dans le secondaire, donnant la solution de l'équation (1). Expliquons-nous.

On dit qu'une équation peut être résolue par une formule si ses racines peuvent être exprimées au moyen d'opétrations arithmétiques, de radicaux, de fonctions exponentielles, logarithmiques, trigonométriques et trigo-

nométriques inverses faisant intervenir les coefficients de cette équation. En ce sens, l'équation du second degré $x^2 + px + q = 0$ admet une formule de résolution de la forme

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \quad (2)$$

L'équation [1]) cubique $x^3 + px + q = 0$ admet aussi une formule de résolution :

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \quad (3)$$

L'utilisation de la formule (3) soulève quelquefois des difficultés et implique la connaissance des nombres complexes.

Il existe une formule de résolution des équations du quatrième degré, mais elle est d'une complexité telle qu'il est préférable de ne pas la citer.

La situation est encore moins brillante pour les équations de degré 5 et supérieur à 5. En 1826, le mathématicien norvégien Abel a démontré que pour $n \geqslant 5$ il n'existe pas de formule donnant la solution de l'équation algébrique

$$a_0x^n + a_1x^{n-1} + \ldots + a_n = 0$$

à l'aide d'opérations arithmétiques et de radicaux. Dans quelques cas particuliers seulement peuvent exister des formules donnant la solution d'équations de degré supérieur à 4 [2]).

Si les mathématiciens s'étaient contentés simplement de résoudre les équations admettant une solution exacte,

---

[1]) Toute équation cubique $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$ peut être ramenée à cette forme moyennant le changement de variable

$$x + \frac{a_1}{3a_0} = y.$$

[2]) Les équations algébriques sont traitées dans:
A. Kurosh « Equations algébriques de degré quelconque », Gostekhizdat, 1951;
I. Chafarévitch « Sur la résolution des équations de degré supérieur », Gostekhizdat, 1954.

c'est-à-dire une solution exprimée par une formule, nous aurions eu droit au dialogue suivant :

L'ingénieur: Mes calculs m'ont conduit à cette équation (et il la montre). Il me faut sa solution d'urgence car je dois achever mon projet de construction dans un mois.

Le mathématicien: Je serais infiniment aisé de vous aider, mais à mon grand regret il n'existe pas de formule pour résoudre votre équation.

L'ingénieur: Ne pouvez-vous pas, vous, établir cette formule?

Le mathématicien: Je n'essayerai même pas, car il est prouvé depuis longtemps qu'il n'existe pas de formule de résolution pour des équations de cette forme.

Il faut croire qu'à l'issue d'un pareil dialogue l'ingénieur se sera fait une piètre opinion du mathématicien et de ses moyens. Mais, par bonheur, ce dialogue n'a pas de raison d'être. Et pour cause : les ingénieurs ont rarement besoin des formules de résolution de telle ou telle équation. Ils se contentent d'une solution approchée, qu'elle ait été déterminée par une formule ou non. Et quand bien même ils viendraient à se servir d'une formule, ce n'est que pour calculer la réponse avec la précision voulue.

Supposons qu'on ait trouvé une formule donnant la réponse sous la forme $x = 3 + \sqrt{13}$. Il est évident qu'on ne peut se servir directement de cette formule (on voit mal comment on pourrait fabriquer une pièce détachée de longueur $3+\sqrt{13}$ cm). Pour les calculs pratiques il faut exprimer $\sqrt{13}$ dans le système décimal avec le degré de précision qui corresponde à la nature du problème à résoudre.

L'ingénieur sera donc pleinement satisfait si le mathématicien lui indique une méthode de calcul des racines d'une équation avec une précision convenable. De nombreuses méthodes de résolution approchée d'une équation ont été élaborées. Certaines d'entre elles font précisément l'objet de cet ouvrage.

§ 2 APPROXIMATIONS SUCCESSIVES

La plupart des méthodes de résolution approchée d'équations s'appuient sur les *approximations successives.* Ces dernières sont utilisées aussi bien pour la résolution d'équations que pour des problèmes d'applications.

La méthode des approximations successives est utilisée en artillerie. S'ils veulent atteindre un but, les servants d'une batterie placent le goniomètre et le viseur dans une position adéquate et tirent un coup. S'ils manquent leur but ils modifient en conséquence l'angle de hausse ou la position du goniomètre ou les deux à la fois, compte tenu des observations faites sur le point d'impact du projectile, puis tirent le coup suivant. Au bout de quelques approximations, le goniomètre et le viseur occupent une position telle que le projectile tiré atteint son but.

Parfois on détermine le point de visée par des approximations successives. Supposons par exemple qu'une batterie de D.C.A. située en O ouvre le feu sur un avion (fig. 1). Si l'on pointe cette arme sur le point A_0 où se trouve l'avion, celui-ci ne sera pas touché: quand le projectile atteindra le point A_0, l'avion occupera une autre position A_1. On détermine aisément ce point si l'on connaît les vitesses du projectile et de l'avion. Cependant, même en pointant les pièces sur le point A_1 l'avion peut ne pas être touché, car l'inclinaison des armes modifie les lois régissant le mouvement de l'obus, et le temps dépensé par ce dernier pour parcourir la distance OA_0 sera différent de celui qu'il mettra pour parcourir la distance OA_1. Cette fois cependant l'erreur de tir sera plus petite. Pour la rendre encore plus faible, il faudrait évaluer le temps que mettrait l'obus pour parcourir la distance OA_1 et localiser la position A_2 de l'avion. Cette position A_2 sera précisément le prochain point de visée. Ensuite, il faut calculer le temps que mettrait l'obus pour atteindre le point A_2 et localiser la posi-

Fig. 1

tion A_3 occupée pendant ce temps par l'avion. Après plusieurs approximations successives on détermine avec la précision voulue le point de visée.

Les approximations successives servent à résoudre de nombreux autres problèmes.

Supposons par exemple que nous ayons à livrer du sable des carrières $A_1, \ldots, A_n$ aux chantiers $B_1, \ldots, B_m$. Supposons par ailleurs que la carrière A_j est d'un rendement de a_j tonnes par jour et que les besoins du chantier B_k sont estimés à b_k tonnes par jour. Supposons, enfin, que le coût du transport d'une tonne de sable de la carrière A_j au chantier B_k se monte à c_{jk} (cette quantité dépend de la distance séparant A_j et B_k, de l'état des routes, etc.).

Le graphique du transport est porté sur le tableau 1; x_{jk} désigne la quantité de sable transportée de la carrière A_j au chantier B_k.

Il est clair que les x_{jk} doivent vérifier les relations suivantes:

$$x_{j1} + x_{j2} + \ldots + x_{jm} \leqslant a_j$$

(on ne peut transporter plus de a_j tonnes de sable par jour

Tableau 1

	B_1	B_2	$\dots$	B_m
A_1	x_{11}	x_{12}	$\dots$	x_{1m}
A_2	x_{21}	x_{22}	$\dots$	x_{2m}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_n	x_{n1}	x_{n2}	$\dots$	x_{nm}

de la carrière A_j)

$$x_{1k}+x_{2k}+\dots+x_{nk}=b_k$$

(les besions du chantier B_k sont limités à b_k tonnes par jour).

Le coût du transport s'élève à:

$$\begin{aligned} C = c_{11}x_{11} + c_{12}x_{12} + \dots + c_{1n}x_{1n} + \\ + c_{21}x_{21} + c_{22}x_{22} + \dots + c_{2n}x_{2n} + \\ \dots\dots\dots\dots\dots\dots\dots \\ + c_{m1}x_{m1} + c_{m2}x_{m2} + \dots + c_{mn}x_{mn}. \end{aligned} \qquad (4)$$

Le graphique doit être conçu de manière à minimiser C. On dresse au départ un plan « grossier » qui est ensuite affiné. On peut par exemple procéder de la manière suivante:

On rattache à la carrière A_1 le plus proche chantier. Si cette carrière a un rendement supérieur aux besoins de ce chantier, on lui rattache le plus proche des chantiers restants. On poursuit cette opération jusqu'à ce que la production de la carrière A_1 soit entièrement répartie. Puis

on recommence cette procédure pour la carrière A_2, A_3, etc. Finalement, chaque chantier se trouve rattaché à une carrière.

Ce schéma ne présente pas néanmoins l'avantage d'être le meilleur, car il peut fort bien rester des chantiers qui soient très éloignés des carrières restantes. Aussi importe-t-il de modifier ce graphique. Certains chantiers sont détachés des carrières affectées d'un petit indice et rattachés à des carrières affectées d'un plus grand indice.

Ces méthodes de modification du graphique dont le but est de diminuer le coût du transport sont étudiées en *programmation linéaire* [1]).

Après plusieurs approximations, on obtient un plan où la somme (4) est minimale ou proche de la minimale.

D'une façon générale, dans l'établissement de tout graphique, horaire, etc., on choisit une approximation grossière qu'on affine progressivement jusqu'à obtenir un résultat satisfaisant.

L'usinage des pièces détachées peut aussi être considéré comme une approximation successive de la forme désirée. On part d'une approximation grossière: une pièce de fonderie ou une autre ébauche. Celle-ci est ensuite façonnée sur une machine-outil où elle acquiert une forme proche de la forme voulue. Elle est ensuite soumise à une machine de précision. Ainsi, de proche en proche, c'est-à-dire après plusieurs approximations, on obtient la pièce voulue.

§ 3 ACHILLE ET LA TORTUE

Les premières approximations successives se rencontrent chez le philosophe grec Zénon d'Elée né entre 490 et 485 avant J.-C. Ce philosophe niait la réalité du mouve-

---

[1]) Cf. A. Solodovnikov « Introduction à l'algèbre et la programmation linéaire », « Prosvéchénié ». M., 1966.

ment en avançant le raisonnement suivant: si Achille, le plus véloce des coureurs grecs, s'aventurait à courir après une tortue il ne la rattraperait jamais. Supposons en effet que 1 000 pas séparent Achille de la tortue et qu'Achille parcourt 10 pas en une seconde que la tortue n'en parcourt qu'un. Dans 100 s Achille aura parcouru 1 000 pas et la tortue 100 pas. Lorsqu'il aura parcouru ces 100 pas en 10 s, la tortue en aura fait 10. Pour couvrir ces 10 pas il lui faudra 1 s, la tortue, elle, aura fait 1 pas. La tortue aura donc toujours de l'avance sur Achille et ce dernier ne pourra jamais la rattraper. Donc, le mouvement n'existe pas.

Le raisonnement de Zénon est un pur sophisme et rien de plus. Le mouvement est le mode d'existence de la matière.

Tout élève calculera sans peine le temps que mettra Achille pour rattraper la tortue. Il lui suffit de résoudre l'équation

$$10x - x = 1000, \tag{5}$$

où x désigne le temps cherché. Il vient

$$x = \frac{1000}{9}\,\text{s} = 111\frac{1}{9}\,\text{s}.$$

Les raisonnements de Zénon peuvent être considérés comme une méthode de résolution approchée de l'équation (5).

En effet, si l'on fait passer x dans le second membre de l'équation (5) et que l'on divise les deux membres de cette équation par 10, il vient

$$x = 100 + \frac{x}{10}. \tag{6}$$

Si l'on néglige le terme $\frac{x}{10}$ (qui est petit devant x), on obtient pour x la première valeur approchée $x_1 = 100$. Pour préciser notre réponse, nous allons porter cette valeur de x_1 dans le deuxième membre de l'équation. Nous obtenons une meilleure approximation pour x, plus exactement $x_2 =$

$= 100 + 10 = 110$. En portant cette valeur dans le deuxième membre de l'équation, nous trouvons l'approximation suivante $x_3 = 100 + \frac{110}{10} = 111$. En poursuivant cette procédure nous obtenons les approximations suivantes:

$$x_1 = 100, \; x_2 = 110, \; x_3 = 111, \; x_4 = 111{,}1, \; \ldots$$

c'est-à-dire les mêmes chiffres que ceux obtenus par Zénon. Ces chiffres sont liés entre eux par la relation

$$x_{n+1} = 100 + \frac{x_n}{10}, \qquad (7)$$

qui permet de les calculer successivement. Lorsque n croît, nous approchons la solution exacte $x = 111\frac{1}{9}$ de l'équation (5).

Soulignons que la méthode adoptée pour résoudre l'équation (5) n'est valable qu'en raison de la petitesse de $\frac{x}{10}$ devant x. Dans le cas contraire, les résultats obtenus n'auraient pas approché la solution exacte. Supposons par exemple qu'Achille ait eu à rattraper non point une tortue, mais une antilope faisant 20 pas à la seconde. Mettons le problème sous forme d'équation:

$$10x - 20x = 1000. \qquad (8)$$

La solution est $x = -100$. Cela veut dire qu'Achille et l'antilope ont couru côte à côte il y a 100 secondes, puis l'antilope a dépassé Achille, ne faisant que creuser l'écart avec le temps.

Essayons maintenant de résoudre l'équation (8) comme l'équation (5). Pour cela faisons passer $20x$ dans le second membre et divisons par 10:

$$x = 100 + 2x. \qquad (9)$$

Posons $x_0 = 0$. Il vient $x_1 = 100$. En portant cette valeur dans le second membre de l'équation (9) nous obtenons la

deuxième approximation $x_2 = 300$. En poursuivant cette procédure nous obtenons les nombres:

$$x_0 = 0,\ x_1 = 100,\ x_2 = 300,\ x_3 = 700,\ \ldots$$

qui, en aucune façon, n'approchent la solution exacte $x = -100$ de l'équation (8).

§ 4 LA DIVISION SUR CALCULATEURS

Le lecteur crédule se demandera probablement pourquoi a-t-il fallu résoudre l'équation (5) par la méthode des approximations successives alors qu'il était si simple d'en trouver la solution exacte. A vrai dire l'équation (5) nous intéresse moins que la méthode de résolution par approximations successives que nous appliquerons dans la suite à des équations plus compliquées.

D'ailleurs la résolution par approximations d'équations de la forme (5) n'a été mise à contribution qu'avec l'apparition des ordinateurs. Certains modèles d'ordinateurs ne peuvent accomplir que trois opérations arithmétiques: l'addition, la soustraction et la multiplication. De plus, ils peuvent diviser par 2^n. Comment expliquer qu'ils effectuent des divisions sur n'importe quel nombre?

Au fait qu'est-ce que la division? Diviser b par a revient à résoudre l'équation $ax = b$. Puisque l'ordinateur peut multiplier et diviser par 2^n, on peut toujours considérer que $1/2 \leqslant a < 1$ (sinon, on multiplie ou divise les deux membres de l'équation $ax = b$ par la puissance respective de 2). Mettons l'équation $ax = b$ sous la forme

$$x = (1 - a)\,x + b. \tag{10}$$

Prenons pour première approximation de x la valeur $x_1 = b$. Soit α_1 l'erreur commise sur cette approximation, c'est-à-

dire $x_1 + \alpha_1 = b/a$. Il vient alors de l'équation (10):

$$x_1 + \alpha_1 = (1 - a)\ (x_1 + \alpha_1) + b =$$
$$= (1 - a)\ x_1 + b + (1 - a)\ \alpha_1. \quad (11)$$

Puisque $1/2 \leqslant a < 1$ on a:

$$0 < 1 - a \leqslant 1/2.$$

Le coefficient $1 - a$ étant relativement petit, nous pouvons dans le second membre de l'équation (11) négliger le terme $(1 - a)\ \alpha_1$ qui est au moins deux fois plus petit que α_1. Nous obtenons:

$$x_1 + \alpha_1 \approx (1 - a)\ x_1 + b.$$

Le nombre

$$x_2 = (1 - a)\ x_1 + b$$

sera notre deuxième approximation.

Soit α_2 l'erreur commise sur l'approximation x_2, c'est-à-dire $x_2 + \alpha_2 = b/a$. De (10) il s'ensuit:

$$x_2 + \alpha_2 = (1 - a)\ x_2 + b + (1 - a)\ \alpha_2.$$

Si l'on néglige le terme $(1-a)\alpha_2$, on obtient l'égalité approchée:

$$x_2 + \alpha_2 \approx (1 - a)\ x_2 + b.$$

Pour troisième approximation on peut donc prendre

$$x_3 = (1 - a)\ x_2 + b.$$

En poursuivant cette procédure on trouve la quatrième approximation:

$$x_4 = (1 - a)\ x_3 + b$$

et ainsi de suite. Les nombres $x_1, x_2, \ldots, x_n, \ldots$ tirés successivement de la relation

$$x_{n+1} = (1 - a)\ x_n + b, \quad (12)$$

s'approchent du nombre b/a. Or, cette formule ne renferme que des additions, soustractions et multiplications, donc elle est utilisable par l'ordinateur.

La méthode de division que nous venons de décrire dérive en réalité de la formule donnant la somme d'une progression géométrique décroissante illimitée. Plus exactement, mettons la fraction b/a sous la forme

$$\frac{b}{a}=\frac{b}{1-(1-a)}.$$

Or, la formule mentionnée entraîne:

$$\frac{b}{1-(1-a)}=b+b(1-a)+b(1-a)^2+\ldots+b(1-a)^{n-1}+\ldots \tag{13}$$

Soit x_n la somme des n premiers termes de cette progression:

$$x_n=b+b(1-a)+\ldots+b(1-a)^{n-1}.$$

Il est évident que

$$x_{n+1}=b+b(1-a)+\ldots+b(1-a)^n=$$
$$=b+(1-a)[b+b(1-a)+\ldots+b(1-a)^{n-1}]=$$
$$=b+(1-a)x_n.$$

Cette formule coïncide avec la formule (12). En remplaçant la fraction b/a par la valeur approchée x_n nous remplaçons donc la somme de la progression illimitée (13) par la somme des n premiers termes. Lorsque le nombre n des termes augmente, cette somme se rapproche de la somme de cette progression (la décroissance de la progression (13) découle du fait que $1/2\leqslant a<1$ et par conséquent $0<1-a\leqslant\leqslant 1/2$).

§ 5 EXTRACTION DE RACINES CARRÉES PAR LA MÉTHODE DES APPROXIMATIONS SUCCESSIVES

Nous allons montrer maintenant comment les approximations successives s'appliquent à l'extraction de racines carrées. La méthode d'extraction de racines carrées enseignée au lycée donne successivement les décimales du nombre cherché. Cette méthode peut également être considérée comme une méthode d'approximation. Elle est toutefois assez complexe et les élèves s'en servent machinalement sans en saisir pleinement le sens. Nous allons décrire une autre méthode d'extraction de racines carrées qui était en usage dans l'Antique Babylone bien avant notre ère. On la trouve dans les travaux d'Héron d'Alexandrie. Elle fut ensuite oubliée jusqu'à l'avènement des ordinateurs, qui s'en servent parfois pour extraire des racines carrées.

Supposons que l'on ait à extraire la racine carrée du nombre 28. Donnons-nous une approximation quelconque de cette racine, par exemple $x_1 = 5$. Soit α_1 l'erreur commise sur cette approximation, c'est-à-dire $\sqrt{28} = 5 + \alpha_1$. Pour déterminer α_1 élevons les deux membres de cette égalité au carré. Il vient

$$28 = 25 + 10\alpha_1 + \alpha_1^2,$$

c'est-à-dire

$$\alpha_1^2 + 10\alpha_1 - 3 = 0. \tag{14}$$

Ainsi pour déterminer α_1 nous obtenons une équation du second degré dont la racine positive est $\alpha_1 = -5 \pm \sqrt{28}$. Pour connaître la valeur exacte de α_1 il nous faut calculer $\sqrt{28}$. De prime abord, nous sommes tombés dans un cercle vicieux : pour connaître $\sqrt{28}$ il faut calculer α_1 et inversement, pour connaître α_1 il faut calculer $\sqrt{28}$.

Usons de l'artifice suivant. L'erreur α_1 commise sur la valeur approchée $x_1 = 5$ n'est pas grande, elle est a

priori inférieure à l'unité. Le nombre α_1^2 est encore plus petit. Nous pouvons donc négliger le terme α_1^2 dans l'équation (14). Pour α_1 nous obtenons ainsi l'équation $10\alpha_1 - 3 \approx 0$, qui entraîne $\alpha_1 \approx 0,3$.

Nous connaissons une valeur approchée de l'erreur α_1. Puisque $\sqrt{28} = 5 + \alpha_1$ nous allons prendre pour deuxième approximation

$$x_2 = 5 + 0,3 = 5,3.$$

Répétons cette procédure en vue de trouver une meilleure approximation pour $\sqrt{28}$. Plus exactement, désignons par α_2 l'erreur commise sur $x_2 = 5,3$, c'est-à-dire posons $\sqrt{28} = x_2 + \alpha_2$. En élevant les deux membres de cette égalité au carré et en négligeant le terme α_2^2, il vient $28 \approx x_2^2 + 2x_2\alpha_2$, d'où:

$$\alpha_2 \approx \frac{28 - x_2^2}{2x_2}.$$

La troisième approximation de $\sqrt{28}$ s'exprime donc par la formule

$$x_3 = x_2 + \frac{28 - x_2^2}{2x_2} = \frac{28 + x_2^2}{2x_2}.$$

Or $x_2 = 5,3$, donc $x_3 = 5,2915...$ De la même façon, en partant de $x_3 = 5,2915$, on trouve l'approximation suivante x_4 par la formule

$$x_4 = \frac{28 + x_3^2}{2x_3} = 5,2915\ldots$$

D'une façon générale si on connaît une valeur approchée x_n de $\sqrt{28}$ l'approximation suivante est donnée par la formule:

$$x_{n+1} = \frac{28 + x_n^2}{2x_n}. \qquad (15)$$

L'approximation est améliorée à chaque pas. On arrête les calculs dès que l'écart entre x_{n+1} et x_n est inférieur à la

précision désirée. Si par exemple il faut extraire $\sqrt{28}$ à 0,0001 près, il suffit de prendre la quatrième approximation et poser $\sqrt{28} = 5,2915$ (en effet $x_3 = 5,2915\ldots$ et $x_4 = 5,2915\ldots$).

On suit la même marche pour extraire la racine carrée de tout nombre positif. Pour calculer $\sqrt{a}$ on choisit une approximation initiale x_1 et l'on détermine l'approximation suivante par la formule:

$$x_{n+1} = \frac{a + x_n^2}{2x_n}. \qquad (16)$$

La formule (16) peut être établie moyennant des raisonnements autres que les précédents. Supposons que nous connaissons la n-ième approximation x_n de $\sqrt{a}$. Comme $\sqrt{a} = \sqrt{x_n \cdot \frac{a}{x_n}}$, on en déduit que $\sqrt{a}$ est moyenne proportionnelle entre x_n et $\frac{a}{x_n}$. Pour approximation de cette moyenne proportionnelle, prenons la moyenne arithmétique des nombres x_n et $\frac{a}{x_n}$ c'est-à-dire posons

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = \frac{x_n^2 + a}{2x_n},$$

ce qui n'est autre que la formule (16).

La méthode d'extraction en approximations des racines carrées décrite ci-dessus consiste à remplacer à chaque pas la moyenne proportionnelle entre les nombres x_n et $\frac{a}{x_n}$ par leur moyenne arithmétique.

Voyons maintenant si l'extraction de racines carrées en approximation connaît le même final que la course d'Achille et de la tortue (on dit alors que l'approximation est convergente) ou d'Achille et de l'antilope (on dit que l'approximation est divergente). Nous allons prouver que l'ex-

traction de racines carrées conduit toujours au résultat cherché, c'est-à-dire que le processus est convergent.

Pour cela comparons les erreurs $\alpha_n = \sqrt{a} - x_n$ et $\alpha_{n+1} = \sqrt{a} - x_{n+1}$ commises sur deux approximations successives. D'après la formule (16), l'erreur α_{n+1} peut être mise sous la forme suivante:

$$\alpha_{n+1} = \sqrt{a} - x_{n+1} = \sqrt{a} - \frac{x_n^2 + a}{2x_n} = -\frac{x_n^2 - 2x_n\sqrt{a} + a}{2x_n}.$$

Or,

$$x_n^2 - 2x_n\sqrt{a} + a = (x_n - \sqrt{a})^2 = \alpha_n^2,$$

donc

$$\alpha_{n+1} = -\frac{\alpha_n^2}{2x_n}. \qquad (17)$$

Nous ne considérerons que les approximations positives x_n du nombre $\sqrt{a}$. De l'égalité (17) on déduit que les erreurs $\alpha_2, \alpha_3, \ldots, \alpha_n, \ldots$ sont négatives. En d'autres termes, *toutes les approximations x_n sont, à partir de la seconde, des approximations par excès* [1]; la première approximation x_1 peut être, quant à elle, une approximation par excès ou par défaut.

On démontre aisément à l'aide de la formule (17) que la *valeur absolue de l'erreur commise sur l'approximation x_n diminue au moins de moitié à chaque pas.* En effet, l'égalité (17) peut être mise sous la forme:

$$\alpha_{n+1} = -\frac{\alpha_n}{2x_n}\alpha_n = \frac{x_n - \sqrt{a}}{2x_n}\alpha_n = \left(\frac{1}{2} - \frac{\sqrt{a}}{2x_n}\right)\alpha_n.$$

Donc

$$|\alpha_{n+1}| = \left|\frac{1}{2} - \frac{\sqrt{a}}{2x_n}\right| |\alpha_n|. \qquad (18)$$

[1] Ceci s'explique par le fait que la moyenne arithmétique de deux nombres est toujours plus grande que leur moyenne proportionnelle.

Or, $x_n > 0$, donc

$$\frac{1}{2} - \frac{\sqrt{a}}{2x_n} < \frac{1}{2}.$$

Par ailleurs, on a démontré plus haut que pour $n \geqslant 2$ on a $x_n > \sqrt{a}$, donc

$$\frac{1}{2} - \frac{\sqrt{a}}{2x_n} > 0,$$

d'où

$$\left| \frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right| < \frac{1}{2}. \qquad (19)$$

Des relations (18) et (19) on déduit que

$$|\alpha_{n+1}| < \frac{1}{2} |\alpha_n|,$$

ce qui démontre notre assertion: à chaque pas la valeur absolue de l'erreur diminue au moins de moitié. Il s'ensuit qu'après le deuxième pas la valeur absolue de l'erreur diminue d'au moins quatre fois, et après le troisième, d'au moins huit fois et ainsi de suite. Il est clair que lorsque n croît, la valeur absolue de l'erreur $\alpha_n = \sqrt{a} - x_n$ diminue et tend vers 0. Ce qui signifie que lorsque n croît, x_n tend vers $\sqrt{a}$.

Voyons maintenant comment le choix de la première approximation x_1 peut influer sur le processus. Observons d'abord que ce choix n'a aucun effet sur le résultat final puisque nous avons démontré que quelle que soit l'approximation initiale x_1 les erreurs $\alpha_2, \ldots, \alpha_n, \ldots$ tendent vers 0 lorsque $n \to \infty$. Si donc est donnée la précision des calculs quelle que soit l'approximation initiale x_1, on obtiendra toujours la même $\sqrt{a}$ avec la précision désirée. Même si le choix de l'approximation initiale était mauvais, nous arriverions en fin de compte au bon résultat. A l'issue d'une dizaine de pas la valeur absolue de l'erreur diminue

de plus de 100 fois ($2^{10} = 1024 \approx 1000$) et à l'issue du quarantième pas, d'au moins un billion (10^{12}). Ainsi, dans le calcul de $\sqrt{2}$ si l'on pose $x_1 = 10^6$ alors $\alpha_1 \approx 10^6$, d'où $|\alpha_{40}| < 10^{-6}$. En d'autres termes, au début, l'erreur était de près d'un million et à la fin, sa valeur absolue est devenue inférieure à un millionième.

Et pourtant le choix de l'approximation initiale influe sur la longueur des calculs. Si la première approximation est mauvaise, un grand nombre de pas sont nécessaires avant que l'écart entre x_{n+1} et x_n ne soit inférieur à la précision voulue dans les calculs. Un bon choix de la première approximation a la vertu d'accélérer considérablement les calculs. Aussi, souvent procède-t-on de la manière suivante: la première approximation x_1 est empruntée dans la table des racines carrées, quant à la formule

$$x_2 = \frac{a + x_1^2}{2x_1}, \tag{20}$$

on ne l'emploie que pour préciser le résultat obtenu.

Cette méthode est particulièrement commode, car la vitesse de décroissance de l'erreur augmente considérablement lorsque x_n se rapproche de $\sqrt{a}$. En effet, lorsque nous avons déduit l'inégalité

$$|\alpha_{n+1}| < \frac{1}{2}|\alpha_n|,$$

nous avons remplacé le terme $\left|\frac{1}{2} - \frac{\sqrt{a}}{2x_n}\right|$ par $\frac{1}{2}$ dans la formule (18). Or, si x_n est proche de $\sqrt{a}$, alors la fraction $\frac{1}{2} - \frac{\sqrt{a}}{2x_n}$ est très petite et donc $|\alpha_{n+1}| = \left|\frac{1}{2} - \frac{\sqrt{a}}{2x_n}\right| \cdot |\alpha_n|$ est très inférieur à $\frac{1}{2}|\alpha_n|$.

Précisons l'assertion faite. A cette fin, outre l'erreur absolue $|\alpha_n| = |\sqrt{a} - x_n|$ considérons l'erreur relative β_n de l'approximation x_n, c'est-à-dire le rapport de l'erreur

absolue $|\alpha_n|$ à la valeur exacte de $\sqrt{a}$. Cette erreur est donnée par la formule

$$\beta_n = \frac{|\alpha_n|}{\sqrt{a}} = \left|1 - \frac{x_n}{\sqrt{a}}\right|.$$

De la formule (17) on déduit la formule suivante pour β_{n+1}:

$$\beta_{n+1} = \frac{|\alpha_{n+1}|}{\sqrt{a}} = \frac{|\alpha_n|^2}{2x_n\sqrt{a}}.$$

Puisque $x_n > \sqrt{a}$, on a:

$$\beta_{n+1} < \frac{|\alpha_n|^2}{2(\sqrt{a})^2} = \frac{1}{2}\beta_n^2.$$

Donc les erreurs relatives β_n des valeurs approchées sont liées par la relation

$$\beta_{n+1} < \frac{\beta_n^2}{2}. \tag{21}$$

Si par exemple l'erreur relative de l'approximation x_n est égale à 0,01, elle n'excède pas 0,00005 pour x_{n+1} et 0,0000000013 pour x_{n+2}. Nous remarquons que la précision des approximations croît de plus en plus vite. On démontre que lorsqu'on est suffisamment près de $\sqrt{a}$, à chaque approximation on double au moins le nombre des décimales vraies.

E x e m p l e. Calculer $\sqrt{238}$ à 0,00001 près.

Dans les tables on a $\sqrt{238} = 15,43$. Prenons 15,43 pour x_1 et déterminons x_2 par la formule

$$x_2 = \frac{15,43^2 + 238}{30,86} = 15,42725\ldots$$

Apprécions la précision de ce résultat. Puisque l'erreur com-

mise sur la valeur 15,43 n'excède pas 0,01 on a $\alpha_1 = 0{,}01$ et par conséquent

$$\beta_1 \approx \frac{0{,}01}{15{,}43} < 0{,}001$$

et

$$\beta_2 < \frac{0{,}001^2}{2} = 0{,}0000005.$$

Ce qui signifie que l'erreur absolue sur l'approximation x_2 n'excède pas $15{,}43 \cdot 0{,}0000005 < 0{,}00001$. En d'autres termes, les sept chiffres du nombre $\sqrt{238} = 15{,}42725\ldots$ sont vrais.

Si on veut obtenir quatorze chiffres significatifs on poussera les calculs jusqu'à la troisième approximation. Notons toutefois que cette précision n'est pratiquement jamais exigée.

Penchons-nous en conclusion sur la particularité suivante de la méthode des approximations successives. Lorsqu'on utilise la méthode habituelle d'extraction de racines carrées, une erreur quelconque commise quelque part fausse complètement les calculs. Il en va autrement lorsqu'on utilise la méthode des approximations successives. Supposons en effet que nous ayons commis une erreur et qu'au lieu de x_n on ait obtenu la valeur y_n. Dans ce cas on peut poursuivre les calculs en considérant que y_n est l'approximation initiale de $\sqrt{a}$ puisque nous avons vu plus haut que le choix de cette approximation conduisait invariablement au résultat avec la précision voulue. Le seul inconvénient est qu'il faudra faire quelques approximations supplémentaires.

Grâce à la particularité que nous venons de mentionner on peut calculer les premières approximations avec une précision moindre et ne prendre la précision voulue que dans les dernières approximations. Ceci a pour effet de réduire la durée des calculs.

§ 6 EXTRACTION DE RACINES K-IÈME PAR LA MÉTHODE DES APPROXIMATIONS SUCCESSIVES

La méthode d'extraction de racines carrées décrite plus haut peut être appliquée à l'extraction de racines k-ième. Nous aurons besoin de la formule suivante [1]:

$$(x+\alpha)^k = x^k + kx^{k-1}\alpha + \ldots, \tag{22}$$

les points de suspension désignent les termes en α^2, α^3, etc.

Prouvons cette formule. On sait que

$$(x+\alpha)^2 = x^2 + 2x\alpha + \alpha^2,$$
$$(x+\alpha)^3 = x^3 + 3x^2\alpha + 3x\alpha^2 + \alpha^3,$$

ce que l'on peut écrire sous la forme

$$(x+\alpha)^2 = x^2 + 2x\alpha + \ldots, \tag{23}$$
$$(x+\alpha)^3 = x^3 + 3x^2\alpha + \ldots \tag{24}$$

La formule (22) est donc démontrée pour $k = 2$ et $k = 3$. Multiplions les deux membres de (24) par $x + \alpha$. Il vient

$$(x+\alpha)^4 = (x^3 + 3x^2\alpha + \ldots)(x+\alpha).$$

En effectuant cette expression on obtient un terme en x^4 ne contenant pas α et deux termes $3x^3\alpha$ et $x^3\alpha$ contenant α à la puissance un, pour ce qui est des autres termes, ils contiennent α à la puissance deux et à des puissances supérieures. On peut donc écrire:

$$(x+\alpha)^4 = x^4 + 3x^3\alpha + x^3\alpha + \ldots = x^4 + 4x^3\alpha + \ldots \tag{25}$$

(comme précédemment, les points de suspension désignent les termes contenant α^2, α^3, etc.).

---

[1]) Cette formule découle du binôme de Newton, mais nous ne la supposerons pas connue du lecteur.

La formule (22) est donc démontrée pour $k = 4$. De la même façon on déduit de la formule (25)

$$(x + \alpha)^5 = x^5 + 5x^4\alpha + \ldots \qquad (26)$$

Il est évident que l'on peut prouver la formule (22) pour tout exposant positif entier k.

Revenons maintenant à l'extraction d'une racine k-ième. Supposons connue une approximation quelconque x_1 de la racine cherchée $\sqrt[k]{a}$ et désignons par α_1 l'erreur sur cette approximation, ce que l'on note $x_1 + \alpha_1 = \sqrt[k]{a}$. Alors $(x_1 + \alpha_1)^k = a$. Or, d'après (22) cette égalité peut s'écrire sous la forme suivante:

$$x_1^k + kx_1^{k-1}\alpha_1 + \ldots = a,$$

où les points de suspension désignent les termes renfermant α_1^2, α_1^3, etc.

Si x_1 approche suffisamment $\sqrt[k]{a}$, alors l'erreur α_1 est petite et on peut donc négliger les termes contenant les grandes puissances de α_1. Nous obtenons ainsi l'équation approchée

$$x_1^k + kx_1^{k-1}\alpha_1 \approx a,$$

d'où

$$\alpha_1 \approx \frac{a - x_1^k}{kx_1^{k-1}},$$

et par conséquent, pour approximation suivante on peut prendre le nombre

$$x_2 = x_1 + \frac{a - x_1^k}{kx_1^{k-1}} = \frac{a + (k-1)\,x_1^k}{kx_1^{k-1}}.$$

D'une façon analogue, en partant de l'approximation x_2, on trouve l'approximation suivante:

$$x_3 = \frac{a + (k-1)\,x_2^k}{kx_2^{k-1}}.$$

D'une façon générale sachant une approximation x_n on peut déterminer l'approximation suivante en appliquant la formule

$$x_{n+1}=\frac{a+(k-1)\,x_n^k}{kx_n^{k-1}}. \qquad (27)$$

Comme dans le cas de l'extraction de racines carrées on démontre que la méthode développée converge quel que soit le choix de l'approximation initiale x_1 (pourvu que celle-ci soit positive). En d'autres termes, les nombres $x_1, x_2, \ldots, x_n, \ldots$ approchent $\sqrt[k]{a}$ quel que soit x_1. On pousse les calculs jusqu'à ce que les nombres x_n et x_{n+1} coïncident dans les limites de la précision demandée.

Exemple. Calculer $\sqrt[3]{970}$ avec une précision de 0,001. Pour $k = 3$ la formule d'approximation (27) prend la forme:

$$x_{n+1}=\frac{a+2x_n^3}{3x_n^2}. \qquad (28)$$

Ici $a = 970$. Posons $x_1 = 10$. La formule (28) donne

$$x_2=\frac{970+2\cdot 10^3}{3\cdot 10^2}=\frac{2970}{300}=9{,}900,$$

$$x_3=\frac{970+2\cdot 9{,}9^3}{3\cdot 9{,}9^2}=\frac{2910{,}60}{294{,}03}=9{,}899.$$

Nous constatons que x_2 et x_3 coïncident dans les limites de la précision demandée. Donc à 0,001 près nous avons:

$$\sqrt[3]{970}=9{,}899.$$

§ 7 MÉTHODE DES ITÉRATIONS

Tous les exemples cités précédemment sont des cas particuliers d'une méthode générale de résolution des équations, appelée *méthode des itérations* ou *méthode des approximations successives*. Etudions-la.

L'équation à résoudre $f(x) = 0$ est mise sous la forme

$$x = \varphi(x). \tag{29}$$

On choisit ensuite une approximation initiale x_1 que l'on porte dans le deuxième membre de l'équation (29). La valeur obtenue $x_2 = \varphi(x_1)$ est prise pour deuxième approximation de la solution. D'une façon générale, si on trouve une approximation x_n, l'approximation suivante x_{n+1} est donnée par la formule

$$x_{n+1} = \varphi(x_n).$$

Supposons qu'après quelques approximations on ait $x_n \approx \approx x_{n+1}$ avec la précision demandée. Ce qui entraîne $x_n \approx \approx \varphi(x_n)$ puisque $x_{n+1} = \varphi(x_n)$, c'est-à-dire que x_n est une valeur approchée de la racine de l'équation $x = \varphi(x)$.

Dans le problème d'Achille et la tortue nous avons mis l'équation

$$10x - x = 1000$$

sous la forme

$$x = 100 + \frac{x}{10}$$

et avons cherché les approximations d'après la formule

$$x_{n+1} = 100 + \frac{x_n}{10}.$$

Dans la division sur ordinateur nous avons écrit l'équation

$$ax = b$$

sous la forme

$$x = (1 - a)x + b$$

et avons cherché les approximations d'après la formule

$$x_{n+1} = (1 - a)x_n + b.$$

Enfin, dans l'extraction de la racine k-ième nous avons mis l'équation

$$x^k = a$$

sous la forme

$$x = \frac{a + (k-1)\,x^k}{kx^{k-1}},$$

puis avons cherché les approximations d'après la formule

$$x_{n+1} = \frac{a + (k-1)\,x_n^k}{kx_n^{k-1}}$$

Examinons une équation plus compliquée, résoluble par la méthode des itérations.

E x e m p l e. Résoudre à 0,001 près l'équation suivante

$$10x - 1 - \cos x = 0. \qquad (30)$$

Mettons l'équation (30) sous la forme suivante

$$x = \frac{1 + \cos x}{10}. \qquad (31)$$

Choisissons une approximation initiale, $x_1 = 0$ par exemple, et portons-la dans le second membre de (31). La valeur

$$x_2 = \frac{1 + \cos 0}{10} = 0,2$$

sera prise pour deuxième approximation de la racine cherchée. En substituant x_2 dans le second membre de l'égalité (31), nous obtenons la troisième approximation

$$x_3 = \frac{1 + \cos 0,2}{10} \approx \frac{1 + 0,98}{10} = 0,198.$$

Nous avons plus loin

$$x_4 = \frac{1 + \cos 0,198}{10} \approx 0,198.$$

Remarquons que l'égalité $x_3 = x_4$ est vérifiée avec une précision de 0,001. Puisque $x_4 = \frac{1 + \cos x_3}{10}$, cela veut dire que le nombre $x_3 = 0,198$ est racine de l'équation $x = \frac{1 + \cos x}{10}$ avec une précision de 0,001.

La méthode des itérations appelle plusieurs questions:

1. La suite des nombres $x_1, x_2, \ldots, x_n, \ldots$ obtenue par itération converge-t-elle toujours vers un nombre ξ?

2. Si l'égalité $\lim\limits_{n\to\infty} x_n = \xi$ a lieu, peut-on en déduire que le nombre ξ est racine de l'équation $x = \varphi(x)$?

3. A quelle vitesse les nombres $x_1, \ldots, x_n, \ldots$ approchent-ils la racine ξ de l'équation $x = \varphi(x)$?

Il est plus facile de répondre à la deuxième question. Supposons que les nombres $x_1, \ldots, x_n, \ldots$ tendent vers ξ. Soit l'égalité $x_{n+1} = \varphi(x_n)$ liant entre elles deux approximations successives. Lorsque n croît, le premier membre tend vers ξ et le second vers $\varphi(\xi)$ [1]. Aussi à la limite avons-nous $\xi = \varphi(\xi)$, c'est-à-dire que ξ est bien racine de l'équation $x = \varphi(x)$.

La réponse à la première question est négative. Soit en effet l'équation

$$x = 10^x - 2.$$

Si l'on fait $x_1 = 1$, on a

$$x_2 = 8,\ x_2 = 10^8 - 2, \ldots$$

Lorsque n croît, les nombres x_n croissent sans tendre vers aucune limite. Par ailleurs, si on écrit cette équation sous la forme $x = \lg(x + 2)$, les approximations convergent et après trois approximations nous obtenons $x = 2{,}38$.

Il faut donc remplacer la première question par la suivante:

Comment faut-il définir $\varphi(x)$ pour que la suite $x_1, x_2, \ldots, x_n, \ldots$ soit convergente?

Avant d'y répondre, donnons une interprétation géométrique de la méthode des itérations.

---

[1]) Nous supposons que $\varphi(x)$ est une fonction continue.

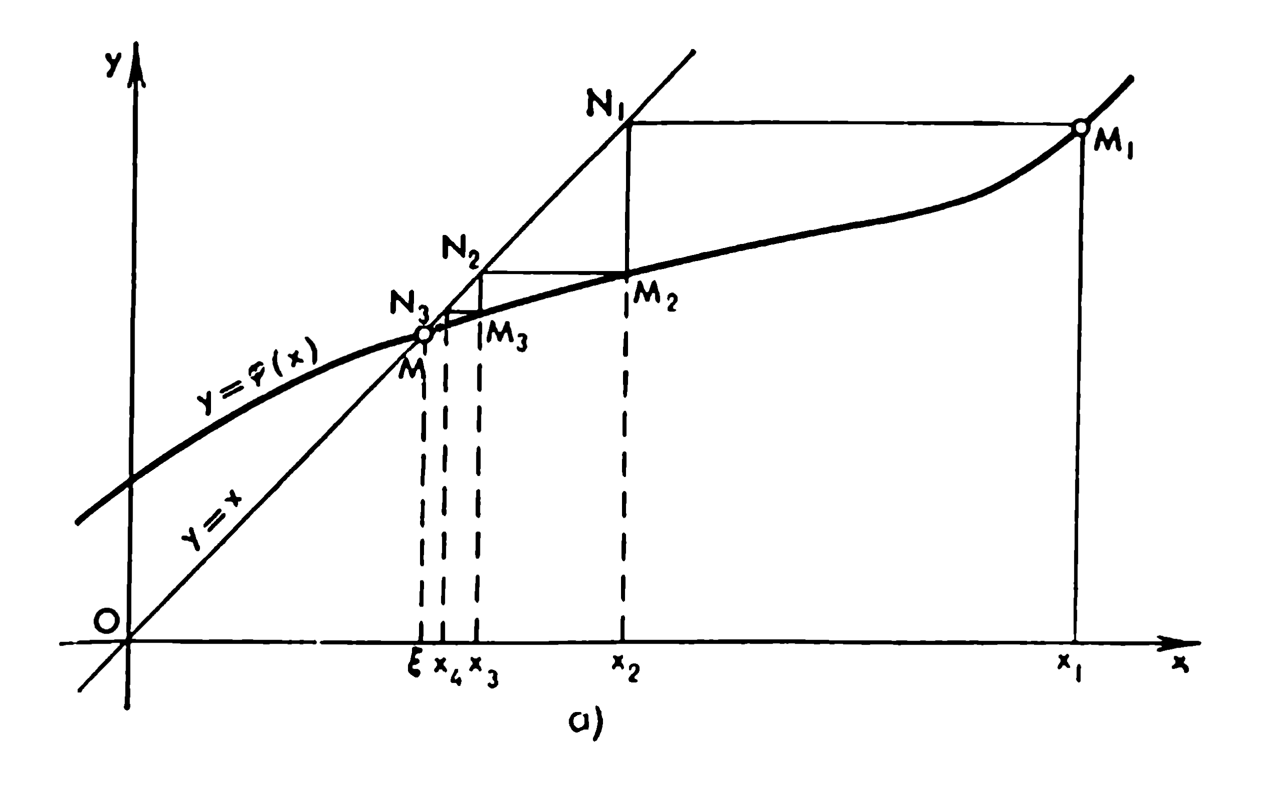
§ 8 INTERPRÉTATION GÉOMÉTRIQUE DE LA MÉTHODE DES ITÉRATIONS

Il est évident que la recherche de la racine ξ de l'équation $x = \varphi(x)$ n'est rien d'autre que la recherche de l'abscisse du point d'intersection M de la courbe $y = \varphi(x)$ avec la droite $y = x$. Supposons donnée une valeur initiale x_1 (fig. 2). Alors le point M_1 de coordonnées $M_1(x_1, \varphi(x_1))$ est situé sur la courbe $y = \varphi(x)$. Menons par ce point une droite parallèle à l'axe des x. Elle coupe la droite $y = x$ au point $N_1(\varphi(x_1), \varphi(x_1))$. Désignons $\varphi(x_1)$ par x_2. Les coordonnées du point N_1 seront alors de la forme $N_1(x_2, x_2)$. Menons ensuite par le point N_1 une droite parallèle à l'axe des y. Elle coupera la courbe $y = \varphi(x)$ au point $M_2(x_2, \varphi(x_2))$. En répétant cette procédure nous obtenons sur la droite $y = x$ le point $N_2(x_3, x_3)$ où $x_3 = \varphi(x_2)$, puis sur la courbe $y = \varphi(x)$ le point $M_3(x_3, \varphi(x_3))$ et ainsi de suite. Si le processus d'approximation converge, les points $M_1, M_2, \ldots M_n, \ldots$ tendent vers le point d'intersection cherché.

Au point de vue géométrique la méthode des approximations successives revient à se déplacer vers le point d'intersection cherché de la courbe avec la droite suivant une ligne brisée dont les sommets sont situés alternativement sur la courbe et sur la droite et dont les côtés sont alternativement parallèles à l'axe des x et à l'axe des y (fig. 2, *a*).

Si la courbe et la droite sont disposées comme sur la fig. 2, *a*, cette ligne brisée rappelle un escalier. Si la courbe et la droite sont disposées comme sur la fig. 2, *b*, la ligne brisée rappelle une spirale.

Le processus d'approximation peut fort bien diverger et n'aboutir à aucun résultat (comme dans le problème d'Achille et l'antilope). Graphiquement, cela signifie que les marches de l'escalier (ou les spires de la spirale) augmentent de dimension et les points $M_1, \ldots, M_n, \ldots$ s'éloignent du point M au lieu de s'en approcher (fig. 3).


y
N1
M1
N2
M2
N3
M3
M
y=φ(x)
y=x
O
ξ
x4
x3
x2
x1
x
a)

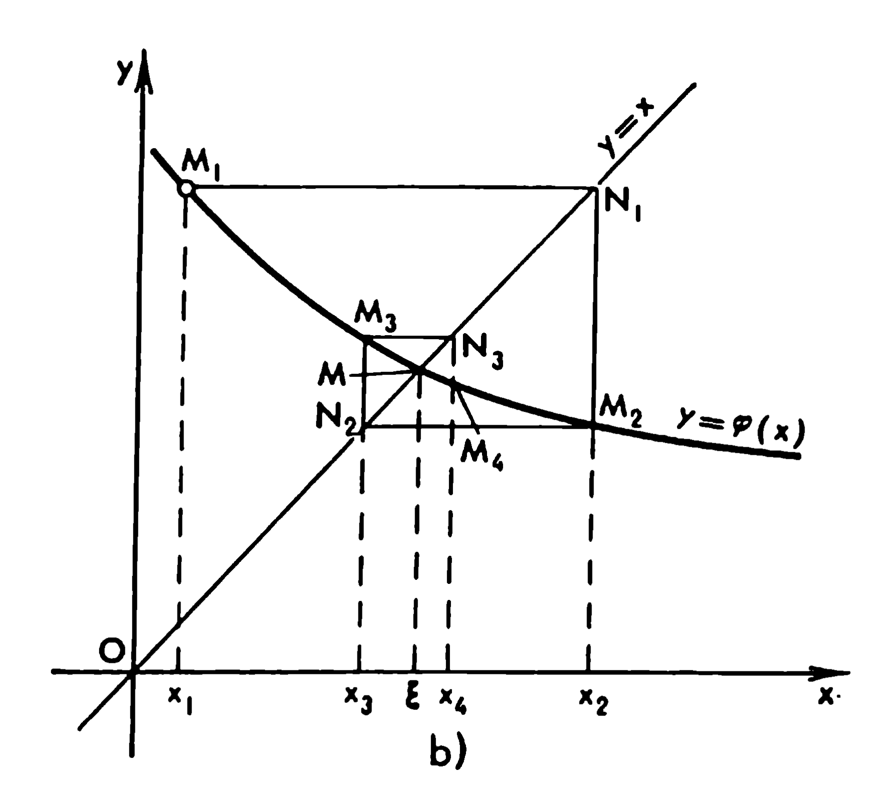
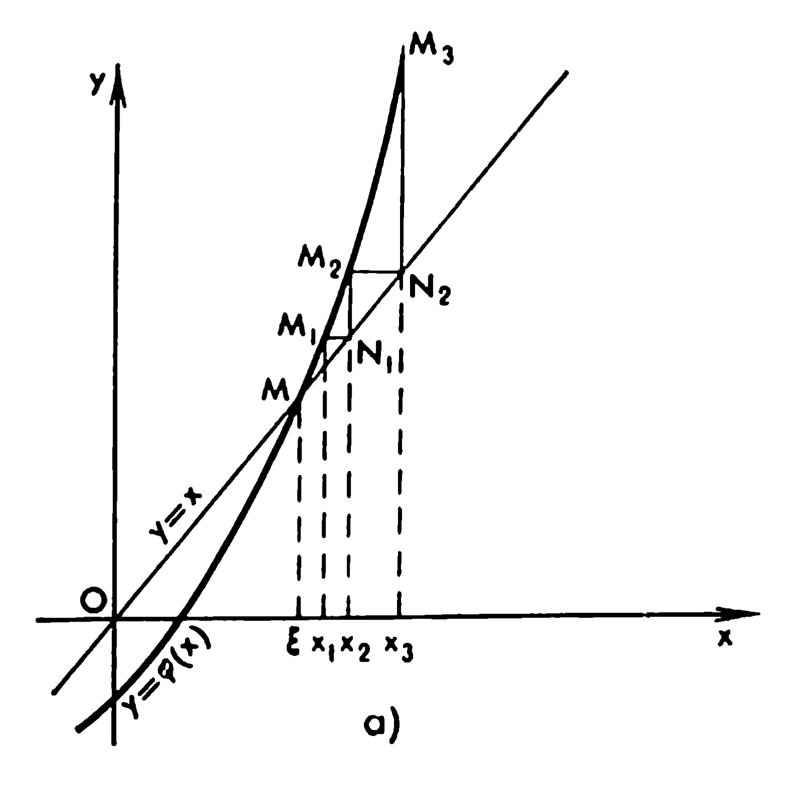

y
y=x
M1
N1
M3
N3
M
N2
M2
M4
y=φ(x)
O
x1
x3
ξ
x4
x2
x
b)

Fig. 2


y
M3
M2
N2
M1
N1
M
y=x
O
ξ x1 x2 x3
x
y=φ(x)
a)

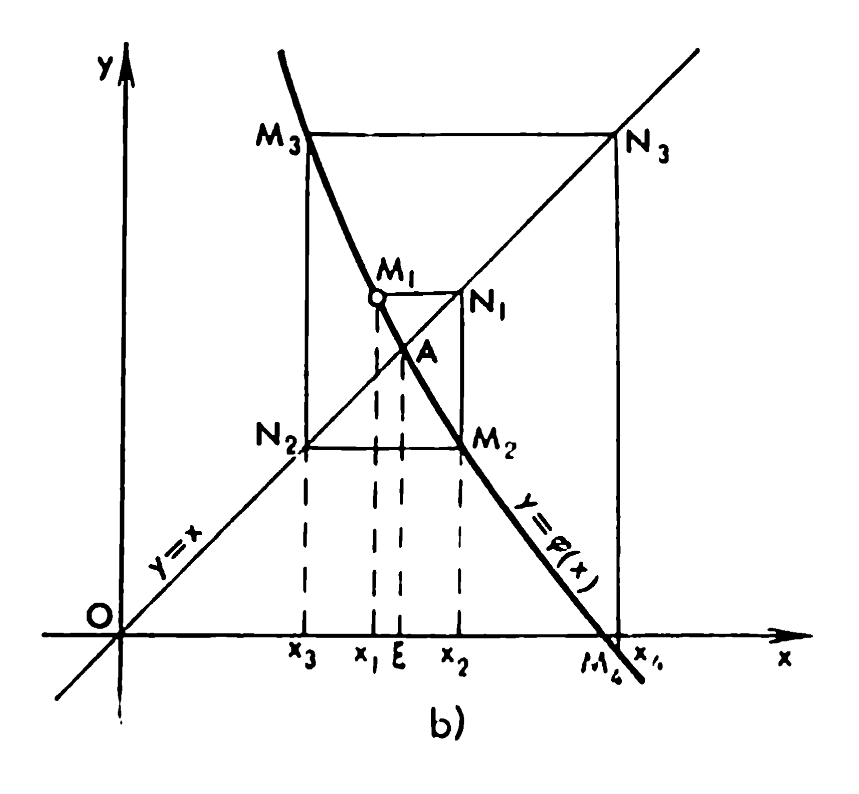

y
M3
N3
M1
N1
A
N2
M2
y=x
y=φ(x)
O
x3 x1 ξ x2
x
b)

Fig. 3

Voyons la différence qui existe entre les figures 2 et 3. Traçons par le point M d'intersection de la droite $y = x$ et de la courbe $y = \varphi(x)$ une droite faisant un angle de 135° avec l'axe des abscisses. Cette droite et la droite $y = x$ divisent le plan en quatre quadrants. Si au voisinage du point M la courbe est située dans les quadrants de gauche et de droite et que l'approximation initiale soit prise dans ce voisinage, le processus itératif converge. Si la courbe est située dans les quadrants supérieur et inférieur, alors le processus itératif est divergent.

Cependant, pour appliquer cette règle il faut préalablement représenter le graphe de la fonction $y = \varphi(x)$, ce qui n'est pas toujours utile. Aussi importe-t-il d'établir un autre critère de convergence, qui permettrait de conclure à la convergence ou la divergence analytiquement, par le calcul et non par une représentation géométrique. Ce critère sera donné au § 10. Auparavant, il nous faudra nous familiariser avec la notion d'application contractante.

§ 9 APPLICATIONS CONTRACTANTES

Soit une fonction $y = \varphi(x)$ définie sur un intervalle $[a, b]$. A tout point x_0 de cet intervalle correspond un point y_0 sur l'axe des ordonnées, plus précisément le point $y_0 = \varphi(x_0)$. Pour construire ce point, on élève une perpendiculaire au point x_0 de l'axe des abscisses, puis au point d'intersection de cette perpendiculaire avec le graphe de la fonction $y = \varphi(x)$ on abaisse une perpendiculaire sur l'axe des ordonnées (fig. 4). La fonction $y = \varphi(x)$ définit donc une application de l'intervalle $[a, b]$ dans l'axe des ordonnées. L'ensemble des points de l'axe des ordonnées correspondant aux points de l'intervalle $[a, b]$ est appelé *image* de cet intervalle. Exemple: l'image de l'intervalle $[2, 5]$ par l'application $y = x^2$ est l'intervalle $[4, 25]$ et celle de l'intervalle $[-1, 6]$ par la même application

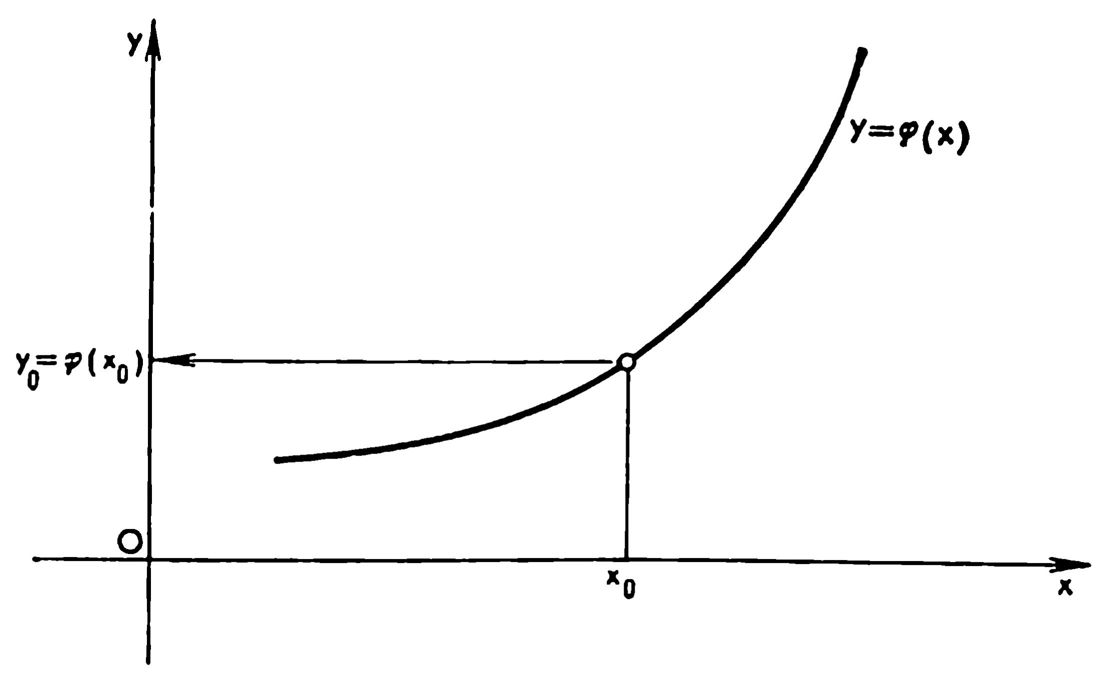


Fig. 4

est l'intervalle [0, 36] (construire le graphe de la fonction $y = x^2$). Il est possible de démontrer que si la fonction $y = \varphi(x)$ est continue sur l'intervalle $[a, b]$, alors l'image de cet intervalle est un intervalle de l'axe des ordonnées. Si, en outre, la fonction $y = \varphi(x)$ est monotone croissante, alors l'image de l'intervalle $[a, b]$ est l'intervalle $[\varphi(a), \varphi(b)]$ et si elle est monotone décroissante, l'intervalle $[\varphi(b), \varphi(a)]$ (fig. 5).

On peut également considérer l'application de l'intervalle $[a, b]$ dans l'axe des abscisses au lieu de l'axe des ordonnées. Pour cela après application dans l'axe des ordonnées on fait pivoter ce dernier de 90° dans le sens des aiguilles d'une montre. Les points de l'intervalle $[a, b]$ se transforment dans des points de l'axe des ordonnées, puis dans des points de l'axe des abscisses. La fonction $\varphi(x)$ définit donc une application de l'intervalle $[a, b]$ dans l'axe des abscisses. Nous noterons cette application: $x \to \varphi(x)$. Si la fonction $\varphi(x)$ est continue, nous obtenons un intervalle de l'axe des abscisses.

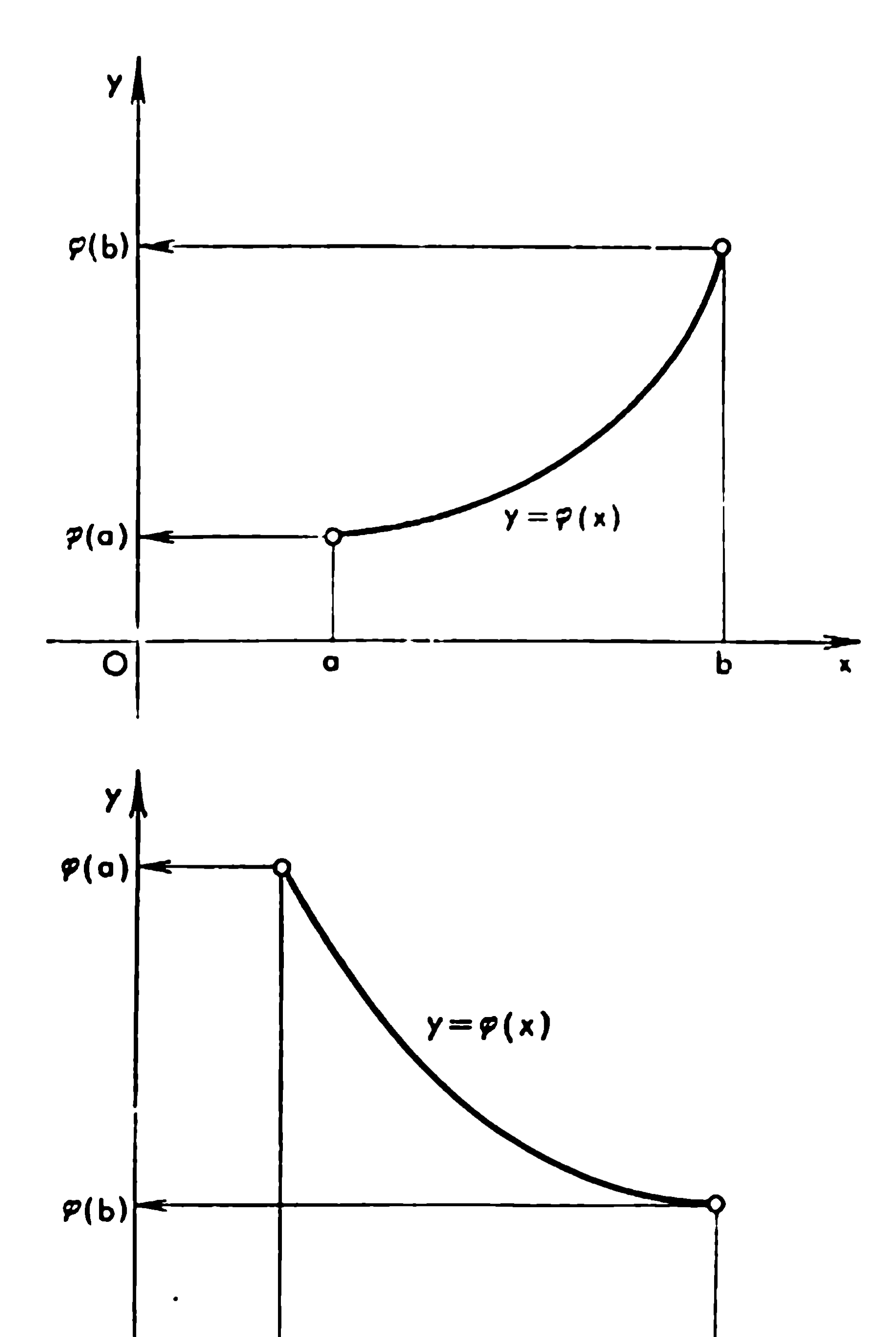


Fig. 5

Il se peut que l'image $[a_1, b_1]$ de l'intervalle $[a, b]$ soit une partie de $[a, b]$. Exemple: $y = \sqrt{x} + 1$ applique $[0, 4]$ sur sa partie $[1, 3]$. Nous dirons dans

ce cas que l'application $\varphi(x)$ associe à l'intervalle $[a, b]$ une de ses parties ou que $\varphi(x)$ est une application de $[a, b]$ sur une de ses parties, soit $[a_1, b_1]$. Si $\varphi(x)$ associe à l'intervalle $[a, b]$ sa partie $[a_1, b_1]$, alors à toute partie de l'intervalle $[a, b]$ elle associe une partie de l'intervalle $[a_1, b_1]$. En particulier, l'intervalle $[a_1, b_1]$ se transforme en sa partie $[a_2, b_2]$ par l'application $\varphi(x)$. D'une façon analogue, l'intervalle $[a_2, b_2]$ se transforme en sa partie $[a_3, b_3]$, etc. Finalement nous obtenons le système d'intervalles

$$[a, b], [a_1, b_1], \ldots, [a_n, b_n], \ldots,$$

dont chacun est une partie de celui qui le précède et tels que $[a_{n+1}, b_{n+1}]$ est l'image de $[a_n, b_n]$ par l'application $\varphi(x)$.

Exemple: l'application $x \to 1 - \frac{1}{x+2}$ associe à l'intervalle [0, 4] sa partie [1/2, 5/6], à l'intervalle [1/2, 5/6] sa partie [3/5, 11/17], et ainsi de suite, chaque intervalle étant une partie du précédent.

Deux cas peuvent se présenter: soit qu'il existe un intervalle $[c, d]$ commun à tous les intervalles $[a_n, b_n]$, soit que ces intervalles n'ont qu'un point commun ξ, et on dit alors que le système d'intervalles $[a_n, b_n]$ *tend vers le point* ξ.

Nous formulerons dans la suite la condition de contraction du système d'intervalles $[a, b]$, $[a_1, b_1]$, ..., $[a_n, b_n]$, ... Introduisons à ces fins la notion importante d'application contractante. Une application $\varphi(x)$ associant à l'intervalle $[a, b]$ sa partie $[a_1, b_1]$ est *contractante* par définition si elle réduit la distance de deux points quelconques de cet intervalle au moins dans le rapport M où $M > 1$. La distance de deux points x_1 et x_2 étant égale à $|x_2 - x_1|$, la condition de contraction s'énonce:

Il existe un nombre q compris entre 0 et 1 tel que quels que soient deux points x_1 et x_2 de l'intervalle $[a, b]$ est véri-

fiée l'inégalité

$$|\varphi(x_2) - \varphi(x_1)| < q\,|x_2 - x_1| \tag{32}$$

(q est égal ici à $1/M$).

La longueur de la partie arbitraire $[c, d]$ de l'intervalle $[a, b]$ est réduite par l'application contractante $\varphi(x)$ d'au moins $M = 1/q$ fois. En effet, supposons que $[c_1, d_1]$ est l'image de l'intervalle $[c, d]$. Alors les points c_1 et d_1 sont respectivement images de deux points quelconques x_1 et x_2 de l'intervalle $[c, d]$,

$$c_1 = \varphi(x_1), \quad d_1 = \varphi(x_2),$$

ce qui entraîne

$$|d_1 - c_1| = |\varphi(x_2) - \varphi(x_1)| \leqslant q\,|x_2 - x_1|.$$

Les points x_1 et x_2 étant situés dans le même intervalle $[c, d]$, la distance $|x_2 - x_1|$ est inférieure à la longueur $|d - c|$ de cet intervalle. Par conséquent,

$$|d_1 - c_1| \leqslant q\,|d - c|.$$

Ce qui démontre notre assertion.

Nous pouvons à présent énoncer la condition de contraction des intervalles $[a_1, b_1], \ldots, [a_n, b_n]$ obtenus à partir de l'intervalle $[a, b]$ par l'application $\varphi(x)$.

Si l'application $\varphi(x)$, *qui à l'intervalle* $[a, b]$ *associe sa partie* $[a_1, b_1]$, *contracte cet intervalle, alors le système d'intervalles* $[a_1, b_1], \ldots, [a_n, b_n]$ *tend vers un point* ξ *de l'intervalle* $[a, b]$.

En effet, l'application $\varphi(x)$ étant contractante, on a, quel que soit n,

$$|b_n - a_n| \leqslant q\,|b_{n-1} - a_{n-1}|.$$

De la même façon

$$|b_{n-1} - a_{n-1}| \leqslant q\,|b_{n-2} - a_{n-2}|.$$

Donc

$$|b_n - a_n| \leqslant q^2\,|b_{n-2} - a_{n-2}|.$$

En reprenant ce raisonnement nous obtenons:

$$| b_n - a_n | \leqslant q^n | b - a |.$$

Comme $0 < q < 1$, la suite des nombres q, q^2, ..., q^n, . . . tend vers 0, la longueur $| b_n - a_n |$ des intervalles $[a_n, b_n]$ tend donc vers 0 lorsque n tend vers l'infini. Or il ne peut exister un intervalle $[c, d]$ commun à tous les intervalles $[a_n, b_n]$. Donc le système d'intervalles

$$[a, b], [a_1, b_1], \ldots, [a_n, b_n], \ldots$$

se réduit à un point.

Etudions en conclusion les applications $\varphi(x)$ telles que l'inégalité (32)

$$| \varphi(x_2) - \varphi(x_1) | < q | x_2 - x_1 |,$$

où $0 < q < 1$ est vérifiée quel que soit le couple de nombres x_1 et x_2. Ces applications sont contractantes sur l'axe numérique tout entier. Montrons que dans ce cas il existe un intervalle qui est contracté par l'application $\varphi(x)$. La relation (23) étant vérifiée quel que soit le couple de nombres x_1 et x_2, il suffit de montrer qu'il existe un intervalle qui se transforme en lui-même par l'application $\varphi(x)$. Prenons un nombre quelconque a et posons $b = \varphi(a)$. Choisissons $q_1 < 1$ tel que $q < q_1$.

Posons

$$R = \frac{| b - a |}{1 - q_1}$$

et montrons qu'à l'intervalle $[a - R, a + R]$ l'application $\varphi(x)$ associe une de ses parties. En effet, soit x un point quelconque de cet intervalle. On a alors $| x - a | < < R$. La relation (32) entraîne

$$| \varphi(x) - b | = | \varphi(x) - \varphi(a) | < q | x - a | \leqslant qR,$$

d'où

$$|\varphi(x)-a|=|\varphi(x)-b+b-a|\leqslant$$
$$\leqslant|\varphi(x)-b|+|b-a|\leqslant qR+|b-a|=$$
$$=qR+(1-q_1)R=(1+q-q_1)R<R.$$

Ce qui montre qu'à tout point de l'intervalle $[a-R, a+R]$ l'application $\varphi(x)$ associe un point de ce même intervalle. Donc, l'application $\varphi(x)$ contracte l'intervalle $[a-R, a+R]$.

§ 10 APPLICATIONS CONTRACTANTES ET MÉTHODE DES ITÉRATIONS

Revenons maintenant à la méthode des itérations. Cette méthode est utilisée pour la résolution d'équations de la forme $x=\varphi(x)$. Si ξ est racine de cette équation, alors $\xi=\varphi(\xi)$ et ce point reste fixe sous l'application $x\rightarrow\varphi(x)$. Donc *résoudre l'équation* $x=\varphi(x)$ *revient à chercher les points fixes* (ou doubles) *de l'application* $\varphi(x)$.

Si l'application $\varphi(x)$ est contractante sur l'intervalle $[a, b]$, il existe toujours un point fixe sur cet intervalle. Pour le prouver, considérons la suite des intervalles

$$[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n], \ldots$$

obtenus de $[a, b]$ en utilisant successivement l'application $\varphi(x)$. L'application $\varphi(x)$ étant contractante, il existe un et un seul point ξ commun à tous les intervalles $[a_n, b_n]$, qui n'est autre qu'un point fixe de l'application $\varphi(x)$.

En effet, à tout intervalle $[a_n, b_n]$ l'application associe sa partie $[a_{n+1}, b_{n+1}]$. Aussi l'image $\varphi(x)$ de tout point de l'ntervalle $[a_n, b_n]$ est-elle comprise dans l'intervalle $[a_{n+1}, b_{n+1}]$ et a fortiori dans l'intervalle $[a_n, b_n]$. Le point ξ étant commun à tous les intervalles $[a_n, b_n]$, son image $\varphi(\xi)$ doit appartenir à tous ces intervalles. Or, le seul point

commun à tous ces intervalles est le point ξ. Donc $\varphi(\xi) = \xi$ et ξ est un point fixe de l'application $\varphi(x)$.

Ainsi, les applications qui contractent l'intervalle $[a, b]$ admettent toujours un point fixe compris dans cet intervalle. Ce point est unique. En effet, s'il existait un autre point fixe η, $\eta = \varphi(\eta)$, on aurait

$$|\eta - \xi| = |\varphi(\eta) - \varphi(\xi)| < q\,|\eta - \xi|.$$

Puisque $0 < q < 1$, cette inégalité n'est vérifiée que si $|\eta - \xi| = 0$, c'est-à-dire si $\eta = \xi$.

Nous pouvons maintenant énoncer la condition suffisante de convergence du processus itératif.

Supposons que la fonction $\varphi(x)$ *définisse une application contractante sur l'intervalle* $[a, b]$. *Alors, quel que soit le point* x_0 *de cet intervalle, la suite* $x_0, x_1, x_2, \ldots, x_n, \ldots$, où $x_{n+1} = \varphi(x_n)$ *converge vers la racine* ξ *de l'équation* $x = \varphi(x)$ *comprise dans cet intervalle.*

Soit en effet $[a_n, b_n]$, $n = 1, 2, \ldots$ une suite d'intervalles déduits successivement de l'intervalle $[a, b]$ par l'application $\varphi(x)$. Le point x_0 étant compris dans l'intervalle $[a, b]$, son image $x_1 = \varphi(x_0)$ est comprise dans l'intervalle $[a_1, b_1]$; l'image $x_2 = \varphi(x_1)$ du point x_1 est comprise dans l'intervalle $[a_2, b_2]$, etc. Donc, quel que soit n, le point x_n est compris dans l'intervalle $[a_n, b_n]$. Les longueurs des intervalles $[a_n, b_n]$ tendant vers 0 lorsque n croît, la suite des points $x_1, \ldots, x_n, \ldots$ tend vers le point commun ξ de ces intervalles.

Pour point initial x_0 on peut donc prendre un point quelconque compris dans l'intervalle $[a, b]$.

Voyons à quelle vitesse les points $x_0, x_1, \ldots, x_n, \ldots$ tendent-ils vers le point ξ. Comme $\xi = \varphi(\xi)$, quel que soit le point c de $[a, b]$, on a:

$$|\varphi(c) - \xi| = |\varphi(c) - \varphi(\xi)| < q\,|c - \xi|. \quad (33)$$

Appliquons l'inégalité (33) aux points $x_0, \ldots, x_n, \ldots$.

Vu que $x_n = \varphi(x_{n-1})$, il vient

$$|x_n - \xi| = |\varphi(x_{n-1}) - \xi| < q\,|x_{n-1} - \xi|;$$

donc quel que soit n, on a

$$|x_n - \xi| < q\,|x_{n-1} - \xi| < q^2\,|x_{n-2} - \xi| < \ldots < q^n\,|x_0 - \xi|.$$

Ainsi *lorsque n croît l'erreur $|x_n - \xi|$ décroît au moins à la vitesse d'une progression géométrique de raison q.*

Illustrons sur des exemples la condition que nous venons de démontrer.

Exemple 1. Peut-on appliquer la méthode des itérations à la résolution de l'équation

$$x = \frac{1}{4+x^2}\,? \qquad (34)$$

Dans ce cas

$$\varphi(x) = \frac{1}{4+x^2}.$$

Quels que soient x_1 et x_2, on a:

$$|\varphi(x_2) - \varphi(x_1)| = \left|\frac{1}{4+x_2^2} - \frac{1}{4+x_1^2}\right| =$$

$$= \frac{|x_1^2 - x_2^2|}{(4+x_2^2)(4+x_1^2)} = \frac{|x_1+x_2|}{|4+x_1^2|\,|4+x_2^2|}\,|x_2 - x_1|.$$

D'après la relation entre la moyenne arithmétique et la moyenne proportionnelle on a

$$|x| = \frac{1}{2}\sqrt{4x^2} \leqslant \frac{4+x^2}{4},$$

donc

$$|x_1 + x_2| \leqslant |x_1| + |x_2| \leqslant \frac{(4+x_1^2)+(4+x_2^2)}{4} =$$

$$= 2 + \frac{x_1^2 + x_2^2}{4} \leqslant 2 + \frac{x_1^2 + x_2^2}{2} + \frac{x_1^2 x_2^2}{8} =$$

$$= \frac{1}{8}(4 + x_1^2)(4 + x_2^2).$$

Nous venons de montrer que quels que soient x_1 et x_2 est vérifiée l'inégalité

$$\frac{x_1+x_2}{(4+x_1^2)(4+x_2^2)} \leqslant \frac{1}{8},$$

d'où

$$|\varphi(x_2)-\varphi(x_1)| \leqslant \frac{1}{8} |x_2-x_1|,$$

ce qui signifie que l'application $\varphi(x)$ est contractante sur l'axe tout entier.

Nous savons déjà que dans ce cas il existe un intervalle qui se transforme en lui-même par cette application. Pour le trouver, posons $a = 0$. L'application $\varphi(x)$ associe au point $a = 0$ le point $b = 1/4$. Par ailleurs, dans le cas considéré $q = 1/8$. Posons $q_1 = 1/4$ et désignons le nombre $\frac{|b-a|}{1-q_1} = \frac{1}{3}$ par R. L'intervalle $[-1/3, 1/3]$ se transforme en lui-même par l'application $\varphi(x)$. Donc cet intervalle contient un point fixe, qui n'est autre que la racine de l'équation (34). Pour trouver ce point, prenons un point quelconque de l'intervalle $[-1/3, 1/3]$, par exemple le point $x_0 = 0$. Appliquons la méthode des itérations. On a :

$$x_1 = \frac{1}{4} = 0{,}25,$$

$$x_2 = \frac{1}{4+0{,}25^2} = \frac{1}{4{,}0625} = 0{,}2461,$$

$$x_3 = \frac{1}{4+0{,}2461^2} = \frac{1}{4{,}0605} = 0{,}2463,$$

$$x_4 = \frac{1}{4+0{,}2463^2} = \frac{1}{4{,}0605} = 0{,}2463.$$

Nous avons $x_3 = x_4$ avec une précision de 0,0001. Ce qui signifie que la racine de l'équation (34), contenue dans l'intervalle $[-1/3, 1/3]$, est égale à 0,2463 avec une précision de 0,0001. L'application $\varphi(x)$ étant contractante sur l'axe tout entier, l'équation (34) ne possède pas d'autre racine.

Exemple 2. Peut-on appliquer la méthode des approximations successives à la résolution de l'équation

$$x = 1 + \sqrt[3]{x}$$

sur l'intervalle $[-1, 8]$?

Ici $\varphi(x) = 1 + \sqrt[3]{x}$. Puisque $\varphi(-1) = 0$ et $\varphi(8) = 3$, nous en déduisons que l'application $\varphi(x)$ transforme l'intervalle $[-1, 8]$ dans lui-même. Cependant, elle n'est pas contractante sur cet intervalle, car, par exemple, pour $x_1 = -0{,}008$ et $x_2 = 0{,}008$ on a

$$|\varphi(x_2) - \varphi(x_1)| = |\sqrt[3]{0{,}008} - \sqrt[3]{-0{,}008}| = 0{,}4 > |x_2 - x_1|.$$

Pour démontrer, dans l'exemple 1, que l'application $\varphi(x)$ était contractante, nous avons utilisé l'inégalité $\sqrt{ab} < \frac{a+b}{2}$. Nous nous proposons maintenant d'établir quelques inégalités qui nous serviront de critère pour montrer si une application est contractante ou non.

Prouvons que pour $x > 0$ est vérifiée l'inégalité

$$\sin x < x < \operatorname{tg} x. \tag{35}$$

Pour cela remarquons que sur la fig. 6 l'aire S_{OAB} du secteur OAB d'angle au centre x est comprise entre les aires des triangles OAB et OAT:

$$S_{\Delta\, OAB} < S_{\text{sect.}\, OAB} < S_{\Delta\, OAT}.$$

Or,

$$S_{\Delta\, OAB} = \frac{R^2 \sin x}{2}, \qquad S_{\Delta\, OAT} = \frac{R^2 \operatorname{tg} x}{2}$$

(R étant le rayon du cercle). L'aire du secteur OAB est égale à $\frac{R^2 x}{2}$ (l'angle est mesuré en radians) ; donc

$$\frac{R^2 \sin x}{2} < \frac{R^2 x}{2} < \frac{R^2 \operatorname{tg} x}{2}.$$

En divisant par $\frac{R^2}{2}$ nous obtenons l'inégalité (35). Celle-ci entraîne

$$x < \arcsin x \text{ pour } 0 < x < 1$$

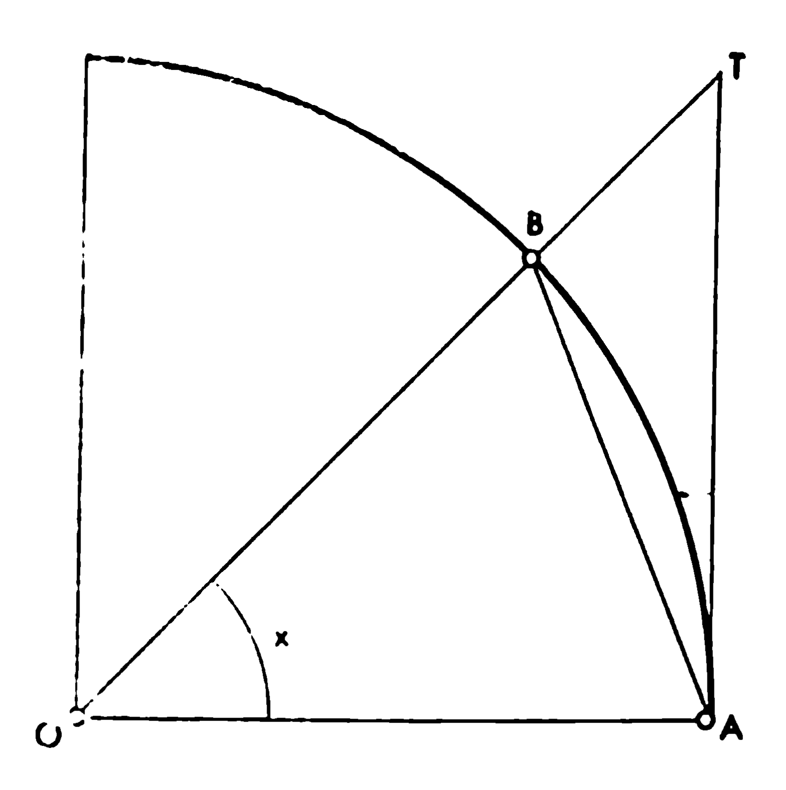


Fig. 6

et

$$x > \operatorname{arctg} x \quad \text{pour} \quad x > 0.$$

Signalons encore deux inégalités:

$$e^x > 1 + x, \quad x > 0,$$
$$\ln(1 + x) < x, \quad 0 < x < 1,$$

dont la démonstration est légèrement plus compliquée.

E x e m p l e 3. Voyons s'il est possible de résoudre l'équation

$$x = 1 + \frac{1}{2} \operatorname{arctg} x \tag{36}$$

par la méthode des itérations. Comme $1 + 1/2 \operatorname{arctg} x > 0$ quel que soit x, l'équation n'admettra que des racines positives. On a

$$\varphi(x) = 1 + \frac{1}{2} \operatorname{arctg} x. \tag{37}$$

D'où

$$|\varphi(x_2) - \varphi(x_1)| = \left| \left(1 + \frac{1}{2} \operatorname{arctg} x_2\right) - \left(1 + \frac{1}{2} \operatorname{arctg} x_1\right) \right| = \\ = \frac{1}{2} |\operatorname{arctg} x_2 - \operatorname{arctg} x_1|.$$

Or, pour $x_1 \geqslant 0$ et $x_2 \geqslant 0$

$$\operatorname{arctg} x_2 - \operatorname{arctg} x_1 = \operatorname{arctg} \frac{x_2 - x_1}{1 + x_1 x_2}$$

et donc

$$| \varphi(x_2) - \varphi(x_1) | =$$

$$= \frac{1}{2} \left| \operatorname{arctg} \frac{x_2 - x_1}{1 + x_1 x_2} \right| < \frac{1}{2} \left| \frac{x_2 - x_1}{1 + x_1 x_2} \right| < \frac{1}{2} | x_2 - x_1 |.$$

L'application (37) est donc contractante sur le demi-axe $[0, \infty)$. Cette application transforme l'intervalle $[0, \sqrt{3}]$ en sa partie $[1, 1 + \pi/6]$. Donc l'équation (36) admet une seule racine comprise dans l'intervalle $[1, 1 + \pi/6]$. Pour la trouver, posons $x_1 = 1$. Il vient

$$x_2 = 1 + \frac{1}{2} \operatorname{arctg} 1 = 1 + \frac{\pi}{8} \simeq 1{,}39,$$

$$x_3 = 1 + \frac{1}{2} \operatorname{arctg} 1{,}39 = 1{,}474,$$

$$x_4 = 1 + \frac{1}{2} \operatorname{arctg} 1{,}474 = 1{,}487,$$

$$x_5 = 1 + \frac{1}{2} \operatorname{arctg} 1{,}487 = 1{,}489,$$

$$x_6 = 1 + \frac{1}{2} \operatorname{arctg} 1{,}489 = 1{,}490,$$

$$x_7 = 1 + \frac{1}{2} \operatorname{arctg} 1{,}490 = 1{,}490.$$

Nous voyons que l'égalité $x_6 = x_7 = 1{,}490$ est vérifiée à 0,001 près. Donc, la racine cherchée est égale à 1,490 à 0,001 près. Puisque l'application $\varphi(x)$ est contractante sur l'axe Ox $(0 \leqslant x < \infty)$, l'équation (36) n'admet pas d'autres racines.

Parfois, lorsque l'équation $x = \varphi(x)$ ne peut être résolue par la méthode des itérations, on la met sous une forme qui permette de lui appliquer cette méthode. Soit par exem-

ple l'équation

$$x = x^3 - 2. \tag{38}$$

Comme

$$\varphi(1) = -1 < 1, \quad \varphi(2) = 6 > 2,$$

cette équation admet une racine comprise dans l'intervalle (1, 2). Or l'application $x^3 - 2$ n'est pas contractante sur cet intervalle puisqu'elle ne le transforme en aucune de ses parties. Mettons l'équation (38) sous la forme

$$x = \sqrt[3]{x+2}.$$

Ici $\psi(x) = \sqrt[3]{x+2}$ et donc

$$|\psi(x_2) - \psi(x_1)| = |\sqrt[3]{x_2+2} - \sqrt[3]{x_1+2}| =$$

$$= \left| \frac{x_2 - x_1}{\sqrt[3]{(x_2+2)^2} + \sqrt[3]{(x_1+2)(x_2+2)} + \sqrt[3]{(x_1+2)^2}} \right|.$$

Dans l'intervalle [1, 2] on a $x_1 \geqslant 1$, $x_2 \geqslant 1$. Donc

$$|\psi(x_2) - \psi(x_1)| \leqslant \frac{1}{3\sqrt[3]{9}} |x_2 - x_1|.$$

Nous avons montré que l'application $\psi(x)$ est contractante sur l'intervalle [1, 2]. Posons $x_1 = 1$ et utilisons la méthode des itérations; nous obtenons

$$x_2 = \sqrt[3]{3} = 1{,}442,$$
$$x_3 = \sqrt[3]{3{,}442} = 1{,}510,$$
$$x_4 = \sqrt[3]{3{,}510} = 1{,}520,$$
$$x_5 = \sqrt[3]{3{,}520} = 1{,}521,$$
$$x_6 = \sqrt[3]{3{,}521} = 1{,}521.$$

Ainsi l'équation (38) admet dans l'intervalle [1, 2] une racine qui est égale à 1,521 avec une précision de 0,001. Elle n'admet pas d'autres racines.

Nous constatons que moyennant une transformation convenable nous avons mis l'équation initiale sous une forme résoluble par la méthode des itérations.

L'indice de convergence de la méthode des itérations n'est pas d'un usage très commode en ce sens qu'il implique la démonstration d'inégalités assez compliquées. Plus bas (§ 21), nous étudierons un corollaire de cet indice qui facilitera grandement la démonstration de la convergence du processus itératif.

§ 11 MÉTHODE DES PARTIES PROPORTIONNELLES

La méthode des itérations fait partie des méthodes les plus générales de résolution approchée des équations. Beaucoup d'autres méthodes de résolution approchée n'en sont que des cas particuliers. Penchons-nous sur l'une d'elles appelée *méthode des parties proportionnelles* ou *d'interpolation linéaire*.

Soit à résoudre l'équation $f(x) = 0$. Ce problème est équivalent à celui qui consiste à rechercher les points en lesquels le graphe de la fonction $y = f(x)$ coupe l'axe des abscisses. Supposons que la fonction $f(x)$ est continue et prend aux points a et b des valeurs de signes contraires. Alors cette fonction s'annule au moins une fois entre les points a et b. En d'autres termes, le graphe de la fonction $y = f(x)$ coupe l'axe des x au moins en un point ξ de $[a, b]$. D'une façon générale, les points de cette nature peuvent être plusieurs (fig. 7). Cependant, si la fonction $y = f(x)$ est m o n o t o n e sur l'intervalle $[a, b]$ et qu'elle prenne à ses extrémités des valeurs de signes contraires, alors le graphe de cette fonction coupe l'axe des x en un seul point ξ compris entre a et b. Pour trouver approximativement ce point, remplaçons sur l'intervalle $[a, b]$ l'arc de courbe $y = f(x)$ par la corde MN et loca-

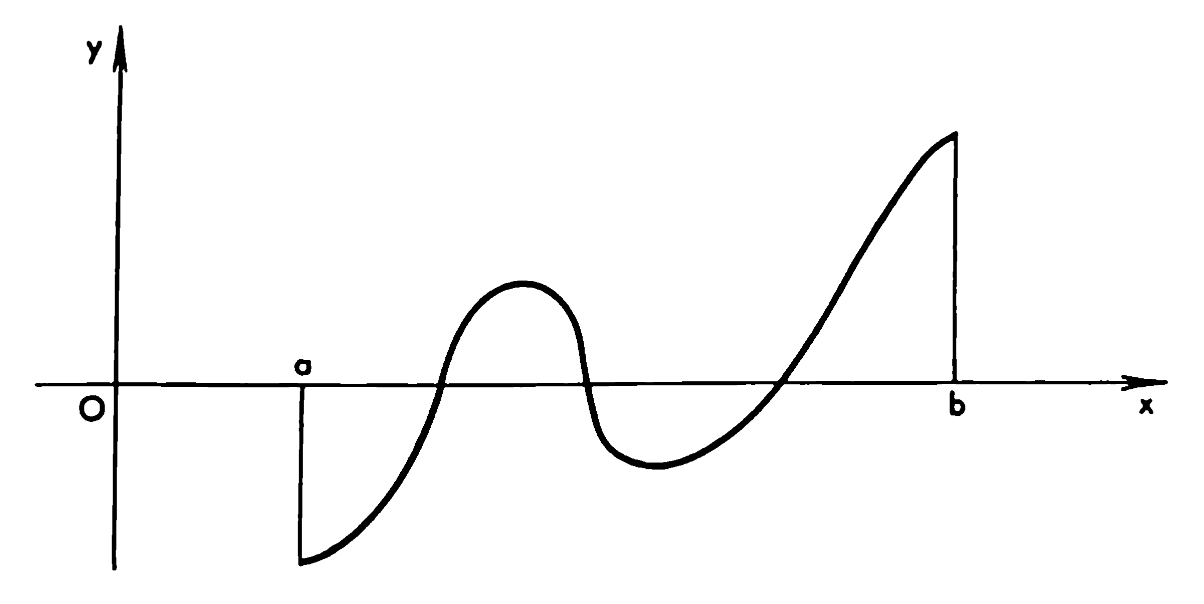


Fig. 7

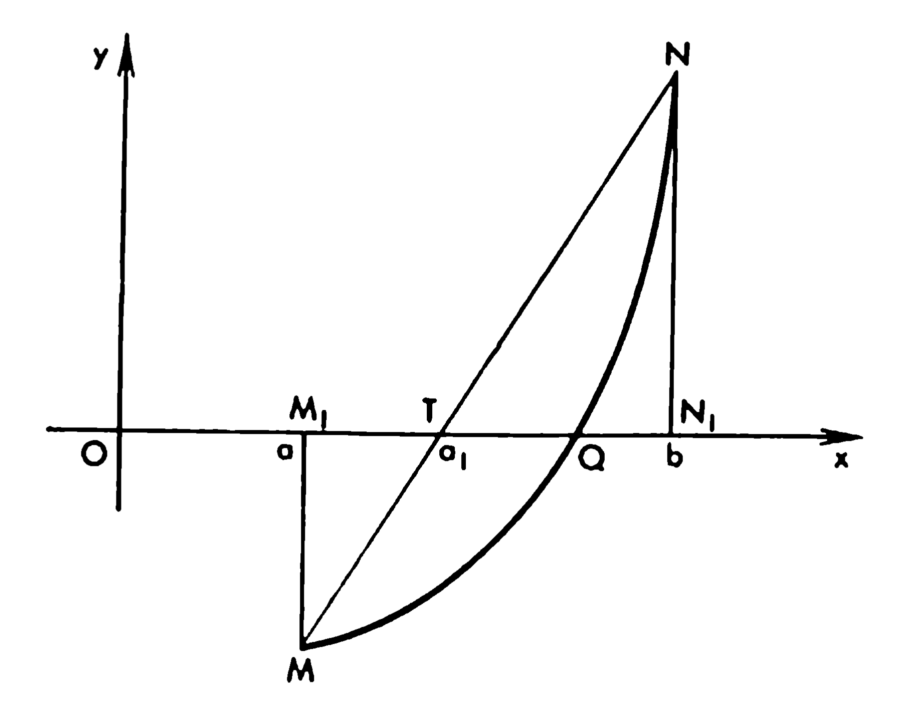


Fig. 8

lisons le point d'intersection de cette corde avec l'axe Ox (fig. 8).

Pour cela considérons les triangles semblables MM_1T et NN_1T. On a $\frac{M_1T}{MM_1} = \frac{TN_1}{N_1N}$. Or d'après la fig. 8, on voit que $M_1T = a_1 - a$, $TN_1 = b - a_1$, $MM_1 = -f(a)$ et $N_1N = f(b)$, où a_1 désigne l'abscisse du point d'intersection de la corde MN avec l'axe Ox. Donc

$$\frac{a_1 - a}{-f(a)} = \frac{b - a_1}{f(b)} .$$

En résolvant cette l'équation on trouve

$$a_1 = \frac{af(b) - bf(a)}{f(b) - f(a)} ,$$

ce que l'on peut mettre sous la forme

$$a_1 = b - f(b) \frac{b - a}{f(b) - f(a)} \tag{39}$$

ou

$$a_1 = a - f(a) \frac{b - a}{f(b) - f(a)} \tag{40}$$

(vérifier ces formules en les réduisant au même dénominateur).

Le nombre a_1 n'est autre que la valeur approchée de la racine de l'équation $f(x) = 0$ comprise entre a et b.

Comme $f(a)$ et $f(b)$ sont de signes contraires, l'un de ces deux nombres est de signe contraire de $f(a_1)$. Si aux points a et a_1 par exemple, la fonction $f(x)$ prend des valeurs de signes contraires, en appliquant la formule (35) à l'intervalle $[a, a_1]$, on obtient l'approximation suivante

$$a_2 = a_1 - f(a_1) \frac{a_1 - a}{f(a_1) - f(a)} \tag{41}$$

pour la racine cherchée. Si, par contre, la fonction $f(x)$ prend des valeurs de signes contraires aux points a_1 et b, en appliquant la formule (40) à l'intervalle $[a_1, b]$ on obtient

$$a_2 = a_1 - f(a_1) \frac{b - a_1}{f(b) - f(a_1)} . \tag{42}$$

Connaissant a_2 on applique la formule (39) à l'intervalle $[a, a_2]$ (respectivement la formule (40) à l'intervalle $[a_2, b]$) et on détermine l'approximation suivante a_3. D'une façon générale, connaissant l'approximation a_n on trouve l'approximation suivante au moyen de la formule

$$a_{n+1} = a_n - f(a_n) \frac{a_n - a}{f(a_n) - f(a)} \tag{43}$$

ou de la formule

$$a_{n+1} = a_n - f(a_n) \frac{b - a_n}{f(b) - f(a_n)} . \tag{44}$$

Voyons maintenant dans quel cas faut-il utiliser l'une ou l'autre de ces formules. Supposons que la courbe est concave vers le haut. Il faut alors joindre les points de la courbe à l'extrémité M ou N où la fonction est positive. Si, au contraire, la courbe est concave vers le bas, il faut joindre les points de cette courbe à l'extrémité où la fonction est négative. Les cas susceptibles de se présenter sont décrits sur la fig. 9. Ces graphes rendent évidente l'assertion suivante:

Soit sur l'intervalle $[a, b]$ *une fonction* $f(x)$ *continue monotone de concavité dirigée dans le même sens et prenant aux extrémités de* $[a, b]$ *des valeurs de signes contraires. Alors, si on choisit convenablement la formule d'approximation, la méthode des cordes définit une suite de points convergeant vers la racine de l'équation* $f(x) = 0$.

Si la formule est mal choisie, le point a_2 peut tomber en dehors de l'intervalle $[a, b]$ comme l'indique la fig. 10.

La méthode décrite est un cas particulier de la méthode des itérations. Supposons que la fonction $f(x)$ ne s'annule pas pour $x = a$. L'équation $f(x) = 0$ est alors équivalente à l'équation

$$x = x - f(x) \frac{x - a}{f(x) - f(a)} . \tag{45}$$

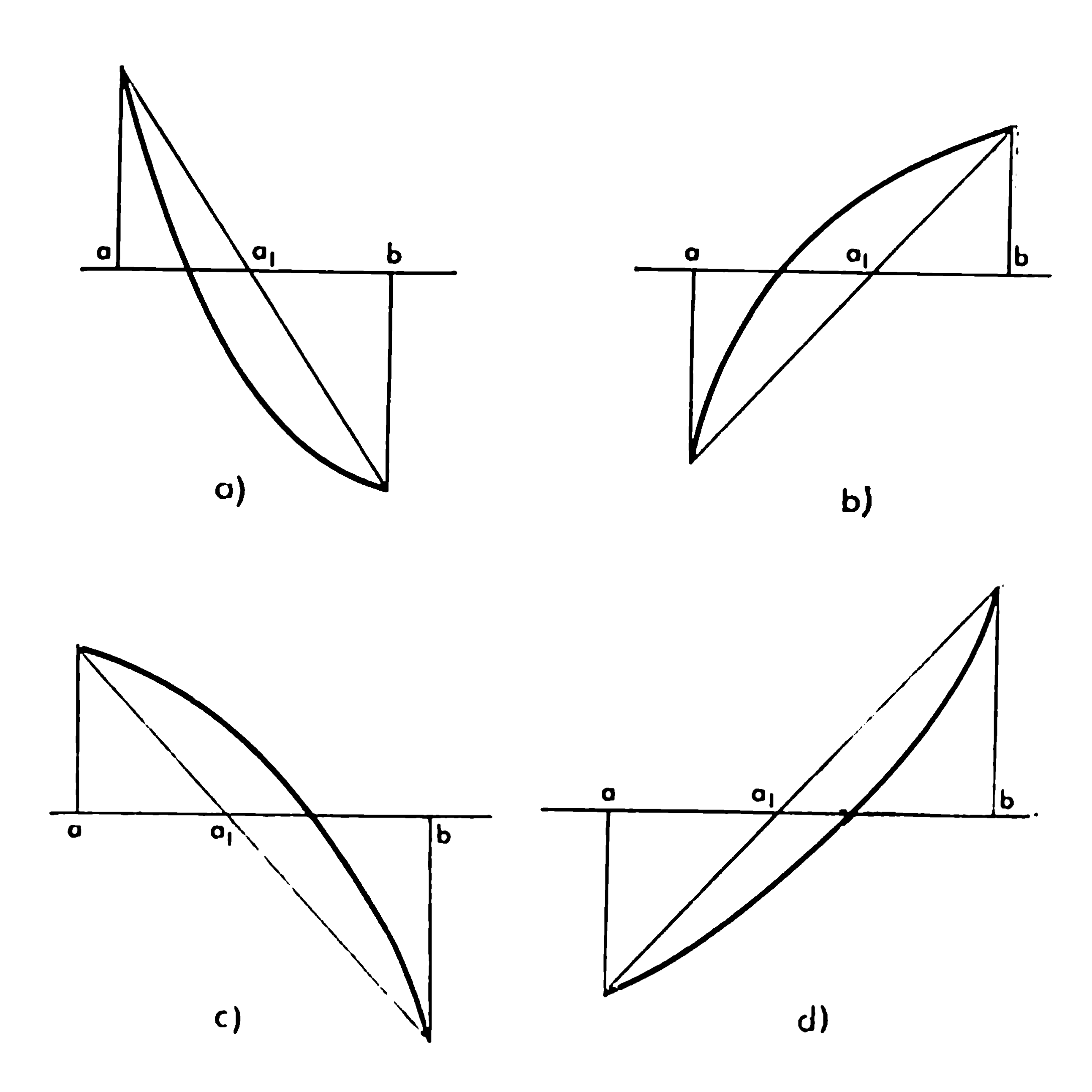


Fig. 9

En effet, si $f(\xi) = 0$, il vient

$$\xi = \xi - f(\xi)\frac{\xi - a}{f(\xi) - f(a)}\,. \tag{46}$$

Inversement, si $\xi \neq a$ et si est vérifiée l'égalité (46), alors $f(\xi) = 0$.

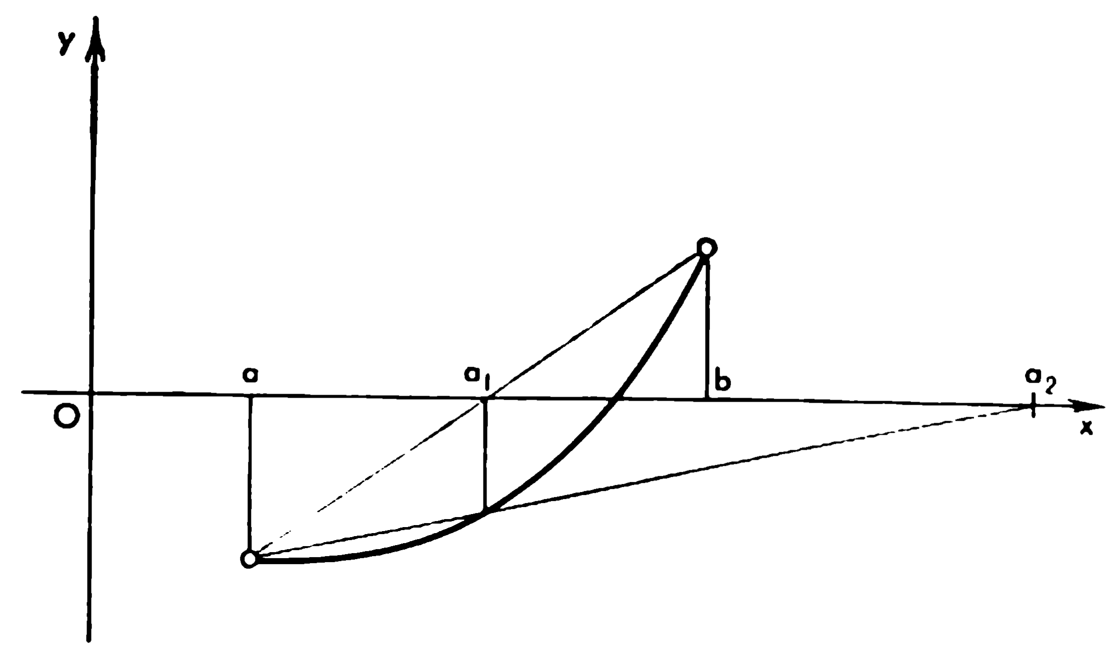


Fig. 10

Or, l'équation (45) est de la forme $x = \varphi(x)$, où

$$\varphi(x) = x - f(x)\frac{x-a}{f(x)-f(a)} = \frac{af(x)-xf(a)}{f(x)-f(a)}.$$

Posons $x_0 = b$ et appliquons la méthode des itérations. Nous obtenons la même suite de nombres $a_1, a_2, \ldots, \ldots, a_n, \ldots$ que pour la méthode des parties proportionnelles:

$$a_{n+1} = a_n - f(a_n)\frac{a_n - a}{f(a_n)-f(a)}.$$

A titre d'exemple appliquons la méthode des parties proportionnelles à la résolution de l'équation

$$x^3 + 3x - 1 = 0. \qquad (47)$$

Ici $f(x) = x^3 + 3x - 1$. Comme $f(0) = -1$ et $f(1) = 3$, l'équation (47) possède au moins une racine dans l'intervalle $[0, 1]$. Si on construit le graphique de la fonction $y = x^3 + 3x - 1$, on constate qu'il est concave vers le haut sur l'intervalle $[0, 1]$. Il convient donc d'utiliser la formule (39). D'après la formule (39), la première appro-

ximation de cette racine est

$$x_1 = b - f(b)\frac{b-a}{f(b)-f(a)} = 1 - 3\cdot\frac{1-0}{3-(-1)} = 0,25.$$

On trouve la seconde approximation par la formule:

$$x_2 = b - f(b)\frac{b-x_1}{f(b)-f(x_1)} = 1 - 3\cdot\frac{1-0,25}{3+0,23} = 0,31.$$

On a ensuite

$$x_3 = 1 - 3\cdot\frac{1-0,31}{3+0,040} = 0,319,$$

$$x_4 = 1 - 3\cdot\frac{1-0,319}{3+0,010} = 0,322,$$

$$x_5 = 1 - 3\cdot\frac{1-0,322}{3+0,0006} = 0,322.$$

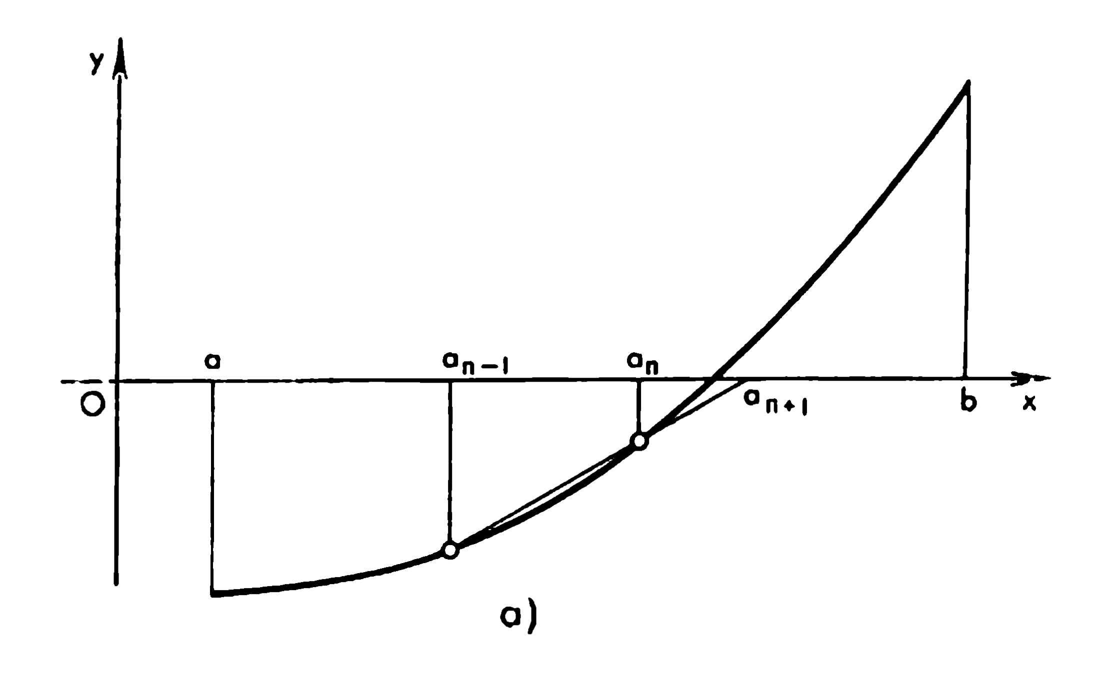
Donc l'équation (47) admet dans l'intervalle [0, 1] une racine égale à 0,322 avec une précision de 0,001.

§ 12 MÉTHODE AMÉLIORÉE DES PARTIES PROPORTIONNELLES

Si la méthode des parties proportionnelles est convergente, la vitesse de convergence sera la même que pour la méthode des itérations: l'erreur sur la racine décroîtra comme une progression géométrique. Il existe une méthode améliorée des parties proportionnelles qui assure une convergence bien plus rapide. Dans la méthode ordinaire à chaque pas nous avons utilisé l'une des extrémités de l'intervalle $[a, b]$ et la dernière approximation obtenue. Au lieu d'agir ainsi nous pouvons prendre les deux dernières approximations, car d'ordinaire elles sont plus proches de la racine cherchée que les extrémités de l'intervalle $[a, b]$.

La formule dans laquelle nous utilisons les deux dernières approximations est de la forme suivante (fig. 11, a):

$$a_{n+1} = a_n - f(a_n)\frac{a_n - a_{n-1}}{f(a_n)-f(a_{n-1})}, \tag{48}$$



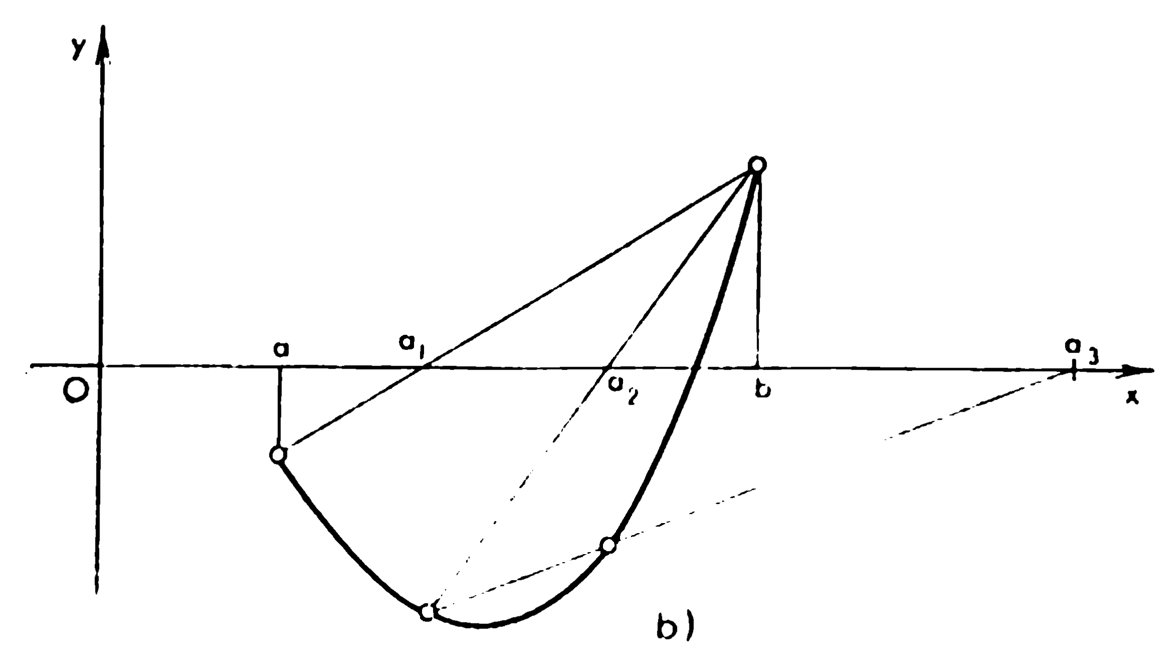


Fig. 11

où a_1 se calcule par la formule (39) et a_2 avec la formule (41) ou (42) selon les signes de $f(a)$, $f(b)$ et $f(a_1)$: si $f(a) < 0$ et $f(b) > 0$, on prend la formule (42) lorsque $f(a_1) < 0$ et la formule (41) lorsque $f(a_1) > 0$.

Si par hasard le point a_3 déterminé à l'aide de la formule (48) est situé en dehors de l'intervalle $[a, b]$, pour le pas suivant il conviendra de prendre au lieu de ce point l'extrémité de $[a, b]$ qui en est la plus proche (fig. 11, b).

Il s'avère que la vitesse de convergence de la méthode améliorée des cordes est bien plus grande que pour la méthode ordinaire. Plus exactement, si ξ est racine de l'équation $f(x) = 0$, on a

$$|a_{n+1} - \xi| < C |a_n - \xi|^t, \qquad (49)$$

où

$$t = \frac{1+\sqrt{5}}{2} \simeq 1{,}618.$$

A titre d'exemple appliquons cette méthode à la résolution de la même équation

$$x^3 + 3x - 1 = 0.$$

Les premières approximations $a_1 = 0{,}25$ et $a_2 = 0{,}31$ sont les mêmes que dans la méthode ordinaire.

L'approximation suivante est donnée par la formule

$$a_3 = a_2 - f(a_2) \frac{a_2 - a_1}{f(a_2) - f(a_1)} =$$

$$= 0{,}31 + 0{,}040 \frac{0{,}31 - 0{,}25}{-0{,}040 + 0{,}234} = 0{,}3223.$$

Nous avons $f(0{,}3223) = 0{,}0004$. Il est clair que $x = 0{,}3223$ est la racine cherchée avec une précision de 0,0001.

§ 13 DÉRIVÉE D'UN POLYNÔME

Dans la résolution de l'équation $f(x) = 0$ par la méthode des itérations beaucoup dépend de la manière dont nous ramenons l'équation donnée à la forme $x = \varphi(x)$. L'une des meilleures méthodes est, dans nombre de cas, celle

proposée par Newton. Cette méthode fait appel à la notion de dérivée. Dans ce paragraphe nous verrons ce qu'est la dérivée d'un polynôme. Ceci nous permettra ensuite d'appliquer la méthode de Newton à la résolution d'équations algébriques, c'est-à-dire d'équations de la forme

$$a_0x^k + a_1x^{k-1} + \ldots + a_k = 0. \tag{50}$$

Soit

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_k$$

un polynôme quelconque. Considérons le polynôme $f(x + \alpha)$, c'est-à-dire l'expression

$$a_0(x+\alpha)^k + a_1(x+\alpha)^{k-1} + \ldots + a_k. \tag{51}$$

Si on développe l'expression (51), on aura des termes ne contenant pas α, des termes en α, en α^2, en α^3, etc. En réduisant les termes semblables, le polynôme $f(x + \alpha)$ prend la forme suivante:

$$f(x+\alpha) = f_0(x) + f_1(x)\alpha + f_2(x)\alpha^2 + \ldots + f_k(x)\alpha^k \tag{52}$$

(comme $f(x)$ est de degré k, la plus grande puissance de α sera égale à k). De plus, il est évident que $f_0(x), \ldots, \ldots, f_k(x)$ sont des polynômes en x.

Exemple. Soit

$$f(x) = 2x^3 - 3x^2 + 6x - 1.$$

Nous avons

$$\begin{aligned} f(x+\alpha) &= 2(x+\alpha)^3 - 3(x+\alpha)^2 + 6(x+\alpha) - 1 = \\ &= 2(x^3 + 3x^2\alpha + 3x\alpha^2 + \alpha^3) - \\ &- 3(x^2 + 2x\alpha + \alpha^2) + 6(x+\alpha) - 1 = \\ &= (2x^3 - 3x^2 + 6x - 1) + (6x^2 - 6x + 6)\alpha + \\ &\qquad + (6x - 3)\alpha^2 + 2\alpha^3, \end{aligned}$$

d'où

$$f_0(x) = 2x^3 - 3x^2 + 6x - 1,$$
$$f_1(x) = 6x^2 - 6x + 6,$$
$$f_2(x) = 6x - 3,$$
$$f_3(x) = 2.$$

On remarquera que le terme $f_0(x)$ coïncide avec $f(x)$. Ce n'est pas un hasard puisque si dans l'égalité (52) on fait $\alpha = 0$ on a $f(x) = f_0(x)$.

Etudions le terme suivant $f_1(x)\alpha$. Le coefficient de α c'est-à-dire le polynôme $f_1(x)$ est appelé *dérivé du polynôme* $f(x)$. Exemple: la dérivée du polynôme $2x^3 - 3x^2 + 6x - 1$ est égale à $6x^2 - 6x + 6$. Il est d'usage de désigner la dérivée de $f(x)$ par $f'(x)$ (on lit f « prime » de x).

Ainsi la *dérivée* $f'(x)$ *du polynôme* $f(x)$ *est le coefficient de* α *dans le développement du polynôme* $f(x+\alpha)$ *suivant les puissances de* α.

En utilisant les notations introduites on peut écrire la formule (52) sous la forme suivante:

$$f(x+\alpha) = f(x) + f'(x)\alpha + \ldots \tag{53}$$

Les points de suspension désignent les termes en $\alpha^2, \alpha^3, \ldots$ $\ldots, \alpha^k$. Par exemple:

$$2(x+\alpha)^3 - 3(x+\alpha)^2 + 6(x+\alpha) - 1 =$$
$$= 2x^3 - 3x^2 + 6x - 1 + (6x^2 - 6x + 6)\alpha + \ldots$$

La notion de dérivée étant introduite, voyons comment calculer cette dernière. Pour cela considérons le polynôme

$$f(x+\alpha) = a_0(x+\alpha)^k + a_1(x+\alpha)^{k-1} + \ldots$$
$$\ldots + a_{k-1}(x+\alpha) + a_k.$$

En développant chaque terme en vertu de la formule $(x+\alpha)^m = x^m + mx^{m-1}\alpha + \ldots$ (cf. § 6) nous obtenons:

$$f(x+\alpha) = a_0(x^k + kx^{k-1}\alpha + \ldots) +$$
$$+ a_1[x^{k-1} + (k-1)x^{k-2}\alpha^2 + \ldots] + \ldots$$
$$\ldots + a_{k-1}(x+\alpha) + a_k =$$
$$= a_0x^k + a_1x^{k-1} + \ldots + a_k +$$
$$+ \alpha[ka_0x^{k-1} + (k-1)a_1x^{k-2} + \ldots + a_{k-1}] + \ldots$$

En comparant cette égalité à l'égalité (53),

$$f(x+\alpha) = f(x) + \alpha f'(x) + \ldots,$$

nous sommes conduits à affirmer ce qui suit:

La dérivée du polynôme

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_{k-1}x + a_k \tag{54}$$

est de la forme

$$f'(x) = ka_0x^{k-1} + (k-1)a_1x^{k-2} + \ldots + a_{k-1}. \tag{55}$$

Par exemple la dérivée du polynôme

$$f(x) = 6x^7 + 8x^3 - 3x^2 - 1$$

est égale à

$$f'(x) = 42x^6 + 24x^2 - 6x.$$

§ 14 MÉTHODE DE NEWTON DE RÉSOLUTION APPROCHÉE DES ÉQUATIONS ALGÉBRIQUES

Revenons à la résolution approchée des équations algébriques. Soit à résoudre l'équation

$$a_0x^k + a_1x^{k-1} + \ldots + a_k = 0. \tag{56}$$

Supposons qu'on connaisse une valeur approchée x_1 d'une de ses racines et montrons comment trouver une valeur plus exacte de cette racine. Soit α_1 l'erreur sur la valeur x_1, c'est-à-dire nous supposons que $x_1 + \alpha_1$ est racine de l'équa-

tion (56). Alors a lieu l'égalité

$$a_0 (x_1 + \alpha_1)^k + a_1 (x_1 + \alpha_1)^{k-1} + \ldots + a_k = 0. \qquad (57)$$

Autrement dit,

$$f(x_1 + \alpha_1) = 0,$$

où $f(x)$ désigne le polynôme

$$a_0 x^k + a_1 x^{k-1} + \ldots + a_k.$$

Or, d'après la formule (31), on a

$$f(x_1 + \alpha_1) = f(x_1) + \alpha_1 f'(x_1) + \ldots,$$

où les points de suspension désignent les termes en $\alpha_1^2, \ldots$ $\ldots, \alpha_1^k$. Donc on tirera α_1 de l'équation

$$f(x_1 + \alpha_1) = f(x_1) + \alpha_1 f'(x_1) + \ldots = 0. \qquad (58)$$

Si l'approximation initiale x_1 est suffisamment bonne, l'erreur α_1 commise sur cette approximation est petite. Alors dans (58) les termes représentés par des points de suspension seront petits devant α_1. En les négligeant, on obtient l'équation approchée

$$f(x_1) + \alpha_1 f'(x_1) \approx 0, \qquad (59)$$

d'où l'on tire

$$\alpha_1 \approx -\frac{f(x_1)}{f'(x_1)}. \qquad (60)$$

Donc la valeur améliorée de la racine est donnée par la formule

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}. \qquad (61)$$

On peut au besoin améliorer cette approximation. La troisième approximation de la racine cherchée est donnée par la formule

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

D'une façon générale, connaissant la n-ième approximation x_n de la racine cherchée on trouve l'approximation suivante

par la formule suivante

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} . \qquad (62)$$

Sous sa forme développée cette formule s'écrit :

$$x_{n+1} = x_n - \frac{a_0 x_n^k + a_1 x_n^{k-1} + \ldots + a_{k-1} x_n + a_k}{k a_0 x_n^{k-1} + (k-1) a_1 x_n^{k-2} + \ldots + a_{k-1}} . \qquad (63)$$

Si les valeurs x_n et x_{n+1} coïncident dans les limites de la précision désirée, le processus itératif est achevé et la racine trouvée.

La méthode de résolution d'équations développée ci-dessus est due au célèbre mathématicien anglais Newton.

La méthode de Newton est étroitement liée à la méthode des itérations. Plus exactement, si les fonctions $y = f(x)$ et $y' = f'(x)$ ne possèdent pas de solution commune, l'équation $f(x) = 0$ est équivalente à l'équation

$$x = x - \frac{f(x)}{f'(x)} . \qquad (64)$$

En appliquant à cette équation la méthode des itérations, on obtient une suite des nombres $x_1, x_2, \ldots, x_n$ liés par la relation

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} , \qquad (65)$$

qui est la même que dans la méthode de Newton. Autrement dit, la méthode de Newton consiste à mettre l'équation $f(x) = 0$ sous la forme (64) et à lui appliquer la méthode des itérations.

E x e m p l e. Résoudre par la méthode de Newton l'équation

$$x^3 - 3x - 5 = 0$$

à 0,001 près en prenant pour première approximation $x_1 = 3$.

Comme le polynôme

$$f(x) = x^3 - 3x - 5$$

admet pour dérivée

$$f'(x) = 3x^2 - 3,$$

la formule (62) s'écrit:

$$x_{n+1} = x_n - \frac{x_n^3 - 3x_n - 5}{3x_n^2 - 3}.$$

Donc

$$x_2 = 3 - \frac{27 - 9 - 5}{27 - 3} = 3 - \frac{13}{24} = 2{,}46,$$

$$x_3 = 2{,}46 - \frac{14{,}89 - 7{,}38 - 5}{18{,}16 - 3} = 2{,}46 - 0{,}165 = 2{,}295,$$

$$x_4 = 2{,}295 - \frac{12{,}088 - 6{,}885 - 5}{15{,}801 - 3} = 2{,}295 - 0{,}016 = 2{,}279,$$

$$x_5 = 2{,}279 - \frac{11{,}837 - 6{,}807 - 5}{15{,}582 - 3} = 2{,}279.$$

Nous remarquons que l'égalité

$$x_4 = x_5$$

est vérifiée à 0,001 près. Donc l'équation $x^3 - 3x - 5 = 0$ admet la racine 2,279 à 0,001 près.

La méthode de calcul approché proposée dans le § 6 est un cas particulier de la méthode de Newton. En effet, calculer $\sqrt[k]{a}$ revient à résoudre l'équation

$$x^k - a = 0.$$

Or, la dérivée du polynôme $x^k - a$ est égale à kx^{k-1} et la formule (62) appliquée à l'équation $x^k - a = 0$ donne

$$x_{n+1} = x_n - \frac{x_n^k - a}{kx_n^{k-1}} = \frac{a + (k-1)\,x_n^k}{kx_n^{k-1}}.$$

On reconnaît la formule dont on se sert pour calculer les approximations de $\sqrt[k]{a}$.

Signalons la différence fondamentale existant entre la résolution de l'équation $x^k - a = 0$ et celle de l'équation algébrique générale

$$a_0x^k + a_1x^{k-1} + \ldots + a_k = 0.$$

Pour la première équation le choix de la première approximation x_1 n'était pas essentiel. En effet, quelle que soit la valeur x_1, au bout d'un certain nombre de pas on arrive à déterminer la racine $\sqrt[k]{a}$ avec la précision requise. Il en est autrement pour la deuxième équation. Certaines valeurs initiales conduisent à une racine, d'autres à d'autres racines, d'autres enfin à aucune racine, la suite des nombres $x_1, x_2, \ldots, x_n, \ldots$ calculés au moyen de la formule (62) ne tendant vers aucune limite définie. Cette suite est donc divergente.

§ 15 SIGNIFICATION GÉOMÉTRIQUE DE LA DÉRIVÉE

La méthode de Newton n'a été exposée que pour les équations algébriques. Pour la généraliser à des équations quelconques il faudrait généraliser la notion même de dérivée à des fonctions quelconques. Pour cela voyons la signification géométrique de la dérivée.

Considérons le graphique du polynôme

$$y = a_0x^k + a_1x^{k-1} + \ldots + a_k$$

et soient, sur ce graphique, deux points M et N (fig. 12) respectivement d'abscisses x et $x + \alpha$. Les ordonnées respectives de ces points sont donc

$$f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_k$$

et

$$f(x + \alpha) = a_0(x + \alpha)^k + a_1(x + \alpha)^{k-1} + \ldots + a_k.$$

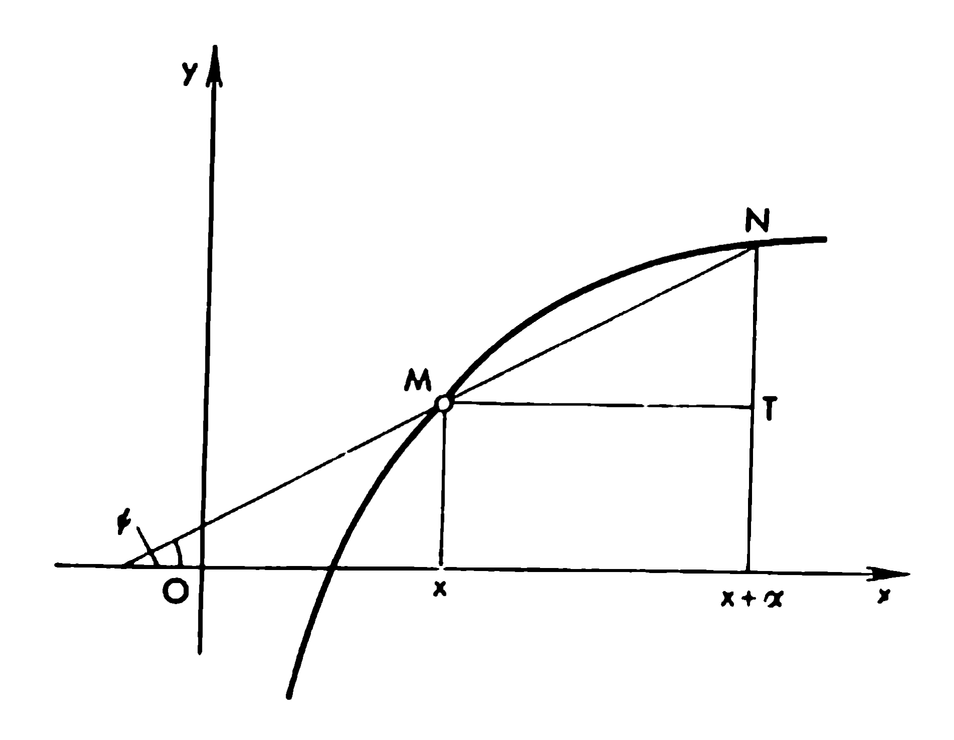


Fig. 12

Joignons les points M et N et calculons le coefficient angulaire k_{sec} [1]) de la sécante MN. On a

$$\operatorname{tg} \psi = \frac{TN}{MT}.$$

Or, le segment MT est égal à la différence des abscisses des points M et N, donc:

$$MT = (x + \alpha) - x = \alpha.$$

Le segment TN est égal lui à la différence des ordonnées de ces points, c'est-à-dire:

$$TN = f(x + \alpha) - f(x).$$

Donc

$$\operatorname{tg} \psi = \frac{TN}{MT} = \frac{f(x+\alpha) - f(x)}{\alpha}.$$

---

[1]) Le *coefficient angulaire* d'une droite est par définition la tangente de l'angle d'inclinaison de cette droite par rapport à l'axe Ox. Si une droite fait un angle de 60° avec l'axe Ox, son coefficient angulaire est égal à $\operatorname{tg} 60° = \sqrt{3}$.

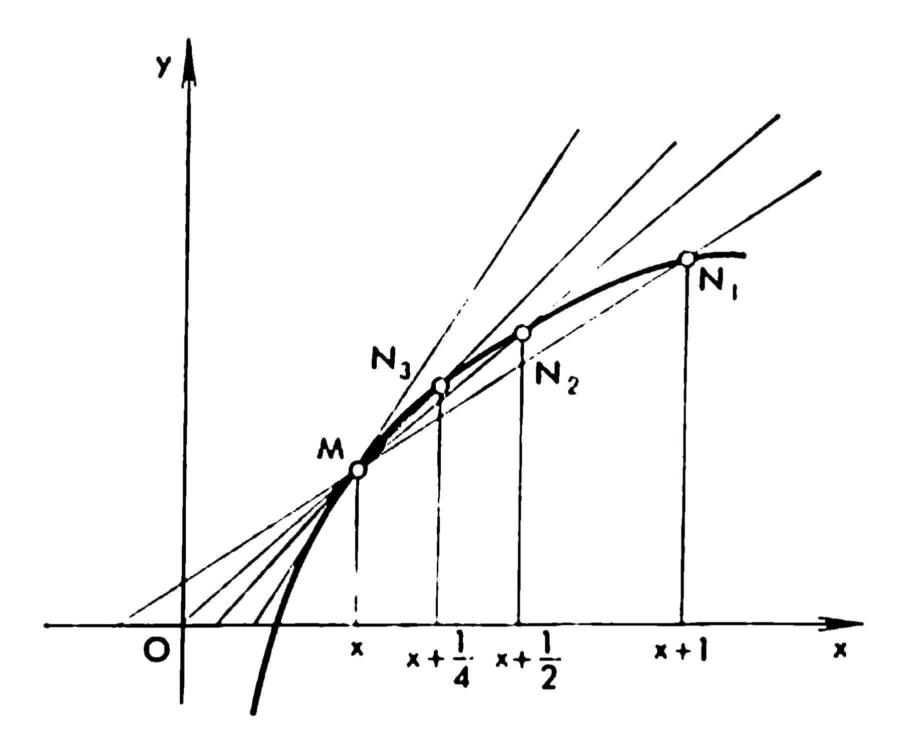


Fig. 13

Or, en vertu de la formule (53), on a

$$f(x+\alpha) = f(x) + \alpha f'(x) + \ldots,$$

où les points de suspension désignent les termes en α^2, α^3, ...Donc

$$\operatorname{tg}\psi = \frac{\alpha f'(x) + \ldots}{\alpha} = f'(x) + \ldots,$$

où cette fois-ci les points de suspension désignent les termes en α, α^2, . . .

Ainsi, le coefficient angulaire de la sécante MN s'exprime au moyen de la formule

$$k_{\text{séc}} = \operatorname{tg}\psi = f'(x) + \ldots \qquad (66)$$

Faisons décroître α. La sécante MN pivotera autour du point M et à la limite, lorsque $\alpha = 0$, elle se transformera en la tangente à la courbe $y = f(x)$ au point M. La fig. 13 représente les positions de la sécante pour $\alpha = 1$; $1/2$; $1/4$.

Or tous les termes représentés dans la formule (66) par des points de suspension s'annulent pour $\alpha = 0$. Donc le coefficient angulaire de la tangente à la courbe $y = f(x)$ au point d'abscisse x est exprimé par la formule

$$k_{\text{tang}} = f'(x). \tag{67}$$

Ainsi, la *dérivée du polynôme $f(x)$ est égale au coefficient angulaire de la tangente à la courbe représentative de ce polynôme au point d'abscisse x.*

E x e m p l e. Trouver l'angle d'inclinaison ψ de la tangente à la courbe

$$f(x) = x^3 - 4x^2 + 5x + 1$$

au point d'abscisse $x = 2$.

Comme on a

$$f'(x) = 3x^2 - 8x + 5,$$

il vient : $f'(2) = 1$. Donc, $\operatorname{tg} \psi = 1$ et $\psi = 45°$.

§ 16 INTERPRÉTATION GÉOMÉTRIQUE DE LA MÉTHODE DE NEWTON

Nous pouvons maintenant donner une interprétation géométrique à la méthode de Newton. Soit à résoudre l'équation $f(x) = 0$ où $f(x)$ est un polynôme quelconque. Géométriquement, ce problème revient à chercher les points d'intersection de la courbe $y = f(x)$ avec l'axe des x, c'est-à-dire les points où $y = 0$.

Supposons que l'on connaisse une valeur approchée x_1 d'une racine de cette équation. Du point N d'abscisse x_1 traçons une tangente à la courbe $y = f(x)$. Si nous faisons un bon choix de l'abscisse x_1, le point d'intersection T de la tangente avec l'axe des x sera plus approché du point d'intersection de la courbe $y = f(x)$ avec l'axe des x que le point M (fig. 14).

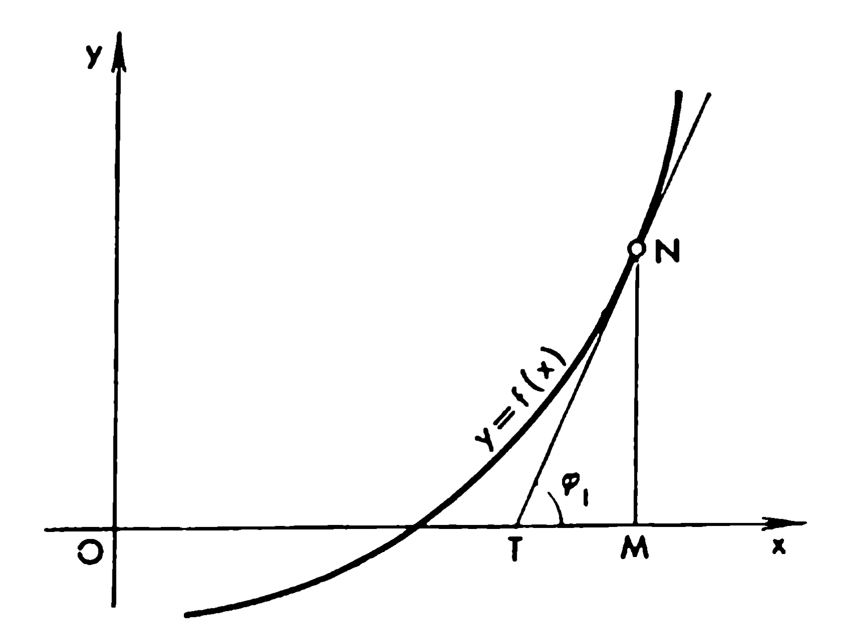


Fig. 14

Pour trouver l'abscisse x_2 du point T, considérons le triangle TMN. Le côté MN de ce triangle n'est autre que la valeur prise par la fonction $y = f(x)$ au point x_1, c'est-à-dire $MN = f(x_1)$. Le côté TM est égal à $x_1 - x_2$. Ce qui entraîne que la tangente de l'angle φ_1 d'inclinaison de la tangente par rapport à l'axe des x est égale à

$$\operatorname{tg} \varphi_1 = \frac{f(x_1)}{x_1 - x_2}, \tag{68}$$

d'où il résulte

$$x_2 = x_1 - \frac{f(x_1)}{\operatorname{tg} \varphi_1}. \tag{69}$$

Or $\operatorname{tg} \varphi_1$ est le coefficient angulaire de la tangente à la courbe $y = f(x)$ menée du point d'abscisse x_1. En vertu donc de la signification géométrique de la dérivée on a $\operatorname{tg} \varphi_1 = f'(x_1)$.

La formule (69) peut donc s'écrire sous la forme

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Ce qui nous donne la deuxième approximation de la racine cherchée. Traçons maintenant la tangente à la courbe

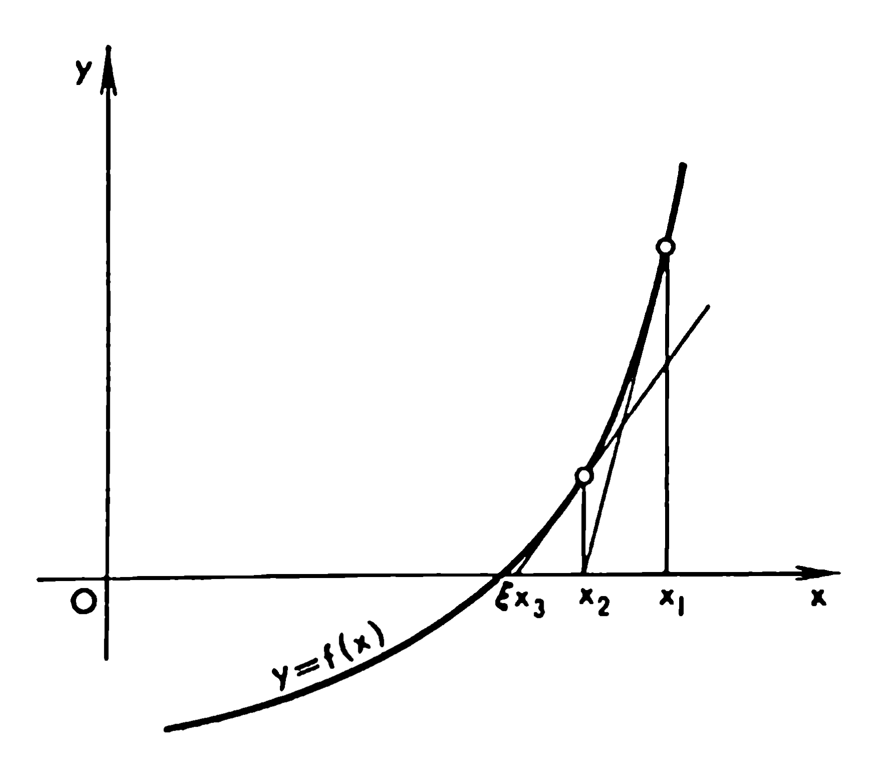


Fig. 15

$y = f(x)$ au point d'abscisse x_2. L'abscisse du point d'intersection de cette tangente avec l'axe des x est définie par la formule

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

D'une façon générale, si on connaît une approximation x_n on trouve l'approximation suivante en construisant la tangente à la courbe $y = f(x)$ au point d'abscisse x_n. L'abscisse du point d'intersection de cette tangente avec l'axe des x donne l'approximation x_{n+1}.

La formule qui permet de calculer x_{n+1} est donc

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad (70)$$

ou encore

$$x_{n+1} = x_n - \frac{f(x_n)}{\operatorname{tg} \varphi_n}, \qquad (70')$$

où φ_n est l'angle d'inclinaison de la tangente à la courbe $y = f(x)$ au point d'abscisse x_n. Cette formule est la même que la formule (62) de la méthode de Newton. Nous venons

de donner ainsi une interprétation géométrique de la méthode de Newton. Elle consiste à remplacer l'arc de la courbe $y = f(x)$ par la tangente à cette courbe. Pour cette raison la méthode de Newton est appelée *méthode des tangentes.*

Sur la fig. 15 on voit comment les points $x_1, x_2, \ldots$ $\ldots, x_n, \ldots$, obtenus par la méthode de Newton, se rapprochent du point d'intersection ξ de la courbe $y = f(x)$ avec l'axe des x.

§ 17 DÉRIVÉES DE FONCTIONS QUELCONQUES

L'interprétation géométrique que nous venons de donner à la méthode de Newton permet de généraliser aisément cette dernière à toute équation de la forme $f(x) = 0$, où $f(x)$ peut ne pas être un polynôme. Plus exactement, pour trouver une solution de cette équation on prend une approximation quelconque x_1 de cette solution. Au point d'abscisse x_1 traçons la tangente à la courbe $y = f(x)$ et désignons par x_2 le point où elle coupe l'axe des x. Du point x_2 on mène une tangente à la courbe $y = f(x)$ et ainsi de suite. Comme pour le cas où $f(x)$ est un polynôme on établit aisément

$$x_{n+1} = x_n - \frac{f(x_n)}{\operatorname{tg} \varphi_n}, \tag{71}$$

où $\operatorname{tg} \varphi_n$ est le coefficient angulaire de la tangente en x_n à la courbe $y = f(x)$.

La formule (71) est encore impropre aux calculs car nous ne savons pas comment définir $\operatorname{tg} \varphi_n$. Il nous faut donc apprendre à calculer les coefficients angulaires des tangentes aux courbes représentatives de toute fonction $y = f(x)$ (et pas seulement pour les graphes des polynômes). Cherchons tout d'abord le coefficient angulaire d'une sécante. Soit M un point arbitraire du graphe de la fonction $y = f(x)$ et MN une sécante. En repre-

nant le même raisonnement que pour les polynômes, on établit que le coefficient angulaire de la sécante MN est donné par la formule

$$k_{\text{sec}} = \operatorname{tg} \psi = \frac{f(x+\alpha) - f(x)}{\alpha}, \qquad (72)$$

où x désigne l'abscisse du point M et $x + \alpha$ celle du point N. Si on fait décroître α, la sécante pivotera autour du point M et tendra vers la tangente à la courbe $y = f(x)$ en ce point (voir fig. 12). On peut donc écrire

$$k_{\text{tang}} = \operatorname{tg} \varphi = \lim_{\alpha \to 0} \frac{f(x+\alpha) - f(x)}{\alpha}. \qquad (73)$$

La limite figurant dans le second membre de cette égalité sera appelée dérivée de la fonction $f(x)$ et désignée par le symbole $f'(x)$, c'est-à-dire qu'on aura

$$f'(x) = \lim_{\alpha \to 0} \frac{f(x+\alpha) - f(x)}{\alpha}. \qquad (74)$$

L'égalité (73) peut maintenant s'écrire sous la forme

$$k_{\text{tang}} = \operatorname{tg} \varphi = f'(x). \qquad (75)$$

Ainsi, quelle que soit la fonction considérée (pas seulement un polynôme), sa dérivée en un point quelconque est égale au coefficient angulaire de la tangente menée à cette courbe en ce point [1]).

Comme $\operatorname{tg} \varphi_n = f'(x_n)$, la formule (71) peut s'écrire sous la forme

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \qquad (76)$$

On reconnaît ici la formule (62). Donc la méthode de Newton a été généralisée à toutes les équations de la forme $f(x) = 0$.

---

[1]) S'il est impossible de tracer une tangente à la courbe $y = f(x)$ au point d'abscisse x, on dit que cette courbe n'admet pas de tangente en ce point.

§ 18 CALCUL DES DÉRIVÉES

Nous avons vu dans le paragraphe précédent que pour trouver le coefficient angulaire de la tangente à la courbe $y = f(x)$ il fallait calculer la limite

$$f'(x) = \lim_{\alpha \to 0} \frac{f(x+\alpha) - f(x)}{\alpha} .$$

En règle générale, ce calcul est souvent compliqué. Cependant, dans nombre de cas importants cette limite est connue. Citons les dérivées des fonctions usuelles:

1. $(a)' = 0.$
2. $(x^k)' = kx^{k-1}.$
3. $(a^x)' = a^x \ln a.$
4. $(\sin ax)' = a \cos ax.$
5. $(\cos ax)' = -a \sin ax.$
6. $(\operatorname{tg} ax)' = \frac{a}{\cos^2 ax} .$
7. $(\operatorname{ctg} ax)' = -\frac{a}{\sin^2 ax} .$
8. $(\log_a x)' = \frac{1}{x \ln a} .$
9. $(\arcsin ax)' = \frac{a}{\sqrt{1-a^2x^2}} .$
10. $(\operatorname{arctg} ax)' = \frac{a}{1+a^2x^2} .$

(Dans les formules 3 et 8 le terme $\ln a$ désigne le logarithme de base $e = 2,71828...$, appelé logarithme naturel ou logarithme népérien.) Signalons que dans la formule 2, l'exposant k peut être un nombre réel quelconque. Par exemple

$$(\sqrt{x})' = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2} x^{\frac{1}{2}-1} = \frac{1}{2\sqrt{x}} ,$$

$$\left(\frac{1}{x^2}\right)' = (x^{-2})' = -2x^{-3} = -\frac{2}{x^3} .$$

Les formules 1-10 ne suffisent pas à calculer les dérivées de toutes les fonctions. Toutefois, si la fonction $f(x)$ se déduit

par des opérations arithmétiques de fonctions dont nous savons calculer les dérivées, on peut facilement établir sa dérivée. Pour cela, on applique les règles suivantes que l'on démontre (comme les formules 1-10 d'ailleurs) en mathématiques supérieures.

1. *La dérivée de la somme de deux fonctions est égale à la somme des dérivées de ces fonctions, c'est-à-dire*

$$[f_1(x)+f_2(x)]'=f_1'(x)+f_2'(x).$$

2. *Si l'un des facteurs d'un produit est constant, on peut faire sortir ce facteur du signe de dérivation*

$$[af(x)]' = af'(x).$$

3. *La dérivée du produit de deux fonctions est fournie par la formule*

$$[f_1(x)\,f_2(x)]'=f_1'(x)\,f_2(x)+f_1(x)\,f_2'(x).$$

4. *La dérivée d'un quotient est fournie par la formule*

$$\left[\frac{f_1(x)}{f_2(x)}\right]'=\frac{f_1'(x)\,f_2(x)-f_1(x)\,f_2'(x)}{[f_2(x)]^2}.$$

La règle de calcul de la dérivée d'un polynôme exposée au § 13 est un corollaire des règles 1 et 2 et de la formule 2.

E x e m p l e 1. Prendre la dérivée du quotient

$$f(x)=\frac{3x^2-x+1}{2x^3+5}.$$

En appliquant la règle 4, il vient

$$f'(x)=\frac{(3x^2-x+1)'\,(2x^3+5)-(3x^2-x+1)\,(2x^3+5)'}{(2x^3+5)^2}.$$

D'autre part, la règle de différentiation d'un polynôme donne

$$(3x^2-x+1)'=6x-1$$

et

$$(2x^3+5)'=6x^2,$$

donc

$$f'(x) = \frac{(2x^3+5)(6x-1)-(3x^2-x+1)\,6x^2}{(2x^3+5)^2} =$$

$$= \frac{-6x^4+4x^3+6x^2+30x-5}{(2x^3+5)^2}.$$

Exemple 2. Prendre la dérivée de la fonction

$$f(x) = \frac{1}{10}\left(\arcsin 3x - \frac{1}{x^2}\right).$$

Solution. D'après les formules 2 et 9 et les règles 1 et 2 on a

$$f'(x) = \frac{1}{10}\,\frac{3}{\sqrt{1-9x^2}} - \frac{1}{10}\left(\frac{-2}{x^3}\right) =$$

$$= \frac{3}{10\sqrt{1-9x^2}} + \frac{1}{5x^3}.$$

Exemple 3. Prendre la dérivée de la fonction

$$f(x) = 10^x \sin 2x.$$

D'après la règle 3 et les formules 3 et 4, on a:

$$f'(x) = (10^x)' \sin 2x + 10^x (\sin 2x)' =$$

$$= 10^x \sin 2x \ln 10 + 10^x \cdot 2 \cos 2x =$$

$$= 10^x (\sin 2x \ln 10 + 2 \cos 2x).$$

Les règles indiquées permettent de prendre les dérivées dans de nombreux cas. Citons encore une règle très importante de dérivation d'une fonction de fonction. Elle s'énonce ainsi:

Si la fonction $y = f(x)$ peut se mettre sous la forme $y = F(z)$ où $z = \varphi(x)$, sa dérivée est fournie par la formule

$$f'(x) = F'(z)\,\varphi'(x), \tag{77}$$

où $z = \varphi(x)$.

Exemple. Prendre la dérivée de la fonction $y = \sin(x^3)$. Cette fonction peut se mettre sous la forme

$y = \sin z$ où $z = x^3$. La dérivée de la fonction $F(z) = \sin z$ est égale à $F'(z) = \cos z$, celle de la fonction $\varphi(x) = x^3$ à $\varphi'(x) = 3x^2$. En appliquant la formule (77) on obtient

$$[\sin(x^3)]' = F'(z)\,\varphi'(x) = \cos z \cdot 3x^2.$$

En remplaçant z par $z = x^3$, on obtient

$$[\sin(x^3)]' = 3x^2 \cos(x^3).$$

§ 19 DÉTERMINATION DES PREMIÈRES APPROXIMATIONS

Voyons maintenant comment choisir les premières approximations. Dans la résolution de l'équation $f(x) = 0$, la première approximation peut être obtenue graphiquement. Pour cela il convient de construire la courbe représentative de la fonction $y = f(x)$ et d'en trouver les points d'intersection avec l'axe des x (aux points d'intersection $y = 0$ et donc $f(x) = 0$).

Si pour une raison ou une autre la représentation graphe de la fonction présente des difficultés (par exemple, si on se sert d'un calculateur), on opte pour une autre méthode de détermination de la première approximation. Pour cela on calcule les valeurs de la fonction pour des valeurs isolées de l'argument (par exemple, pour les valeurs entières de l'argument prises dans un intervalle connu). Si la fonction $y = f(x)$ est continue (c'est-à-dire si son graphique ne comporte pas de discontinuités), et si en a et b elle prend des valeurs de signes contraires (fig. 16, a), alors elle s'annule dans l'intervalle $[a, b]$ (si le graphique de la fonction comporte des discontinuités, celle-ci peut réaliser un saut et passer d'une valeur négative à une valeur positive sans s'annuler (fig. 16, b)). Les valeurs a et b peuvent être prises pour premières approximations de la racine de l'équation $f(x) = 0$.

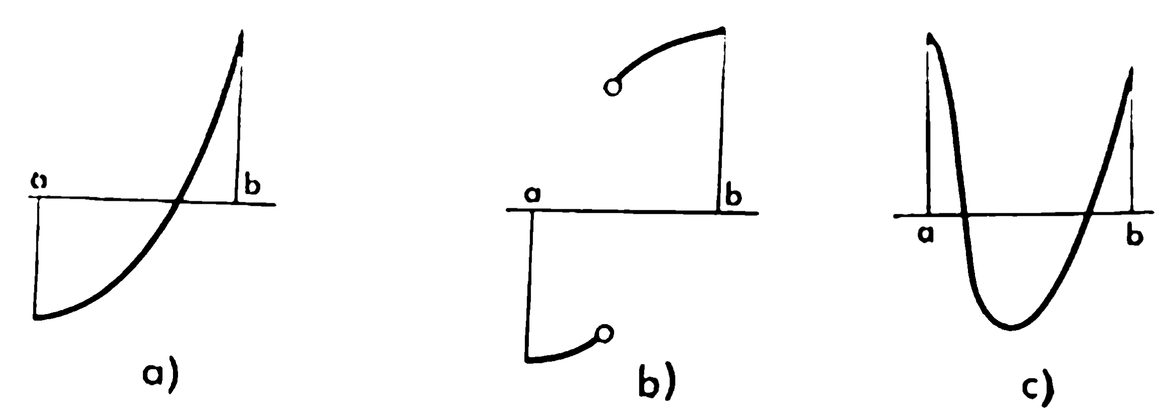


Fig. 16

Il peut se produire aussi que la fonction s'annule plusieurs fois dans l'intervalle $[a, b]$. Ainsi, sur la fig. 16, *c* la fonction $y = f(x)$ prend des valeurs de même signe aux extrémités a et b et s'annule deux fois dans l'intervalle $[a, b]$.

Nous obtenons ainsi deux points : a et b. Pour déterminer lequel d'entre eux conviendra le mieux pour première approximation x dans la méthode de Newton, étudions la fig. 17. On voit sur la fig. 17, *a* et *b* que *si la courbe est concave vers le haut, il convient alors de prendre pour première approximation celui des points a et b où la fonction f (x) est positive. En choisissant une autre approximation, on peut faire en sorte que le point* x_2 *tombe en dehors de l'intervalle* $[a, b]$. *De la même façon, si la courbe est concave vers le bas, il convient de prendre pour première approximation le point où la fonction f (x) est négative* (fig. 17, *c*, *d*).

Cette règle est d'un usage commode si l'on connaît le graphe de la fonction $y = f(x)$. Si la représentation graphique de cette fonction n'est pas connue, des calculs supplémentaires s'imposent pour préciser le sens de sa concavité. Pour cela on fait appel à la dérivée seconde. On appelle *dérivée seconde de la fonction* $f(x)$ la dérivée de sa dérivée. Si, par exemple, est donnée la fonction

$$f(x) = x^3 - 4x^3 + 3x - 1,$$

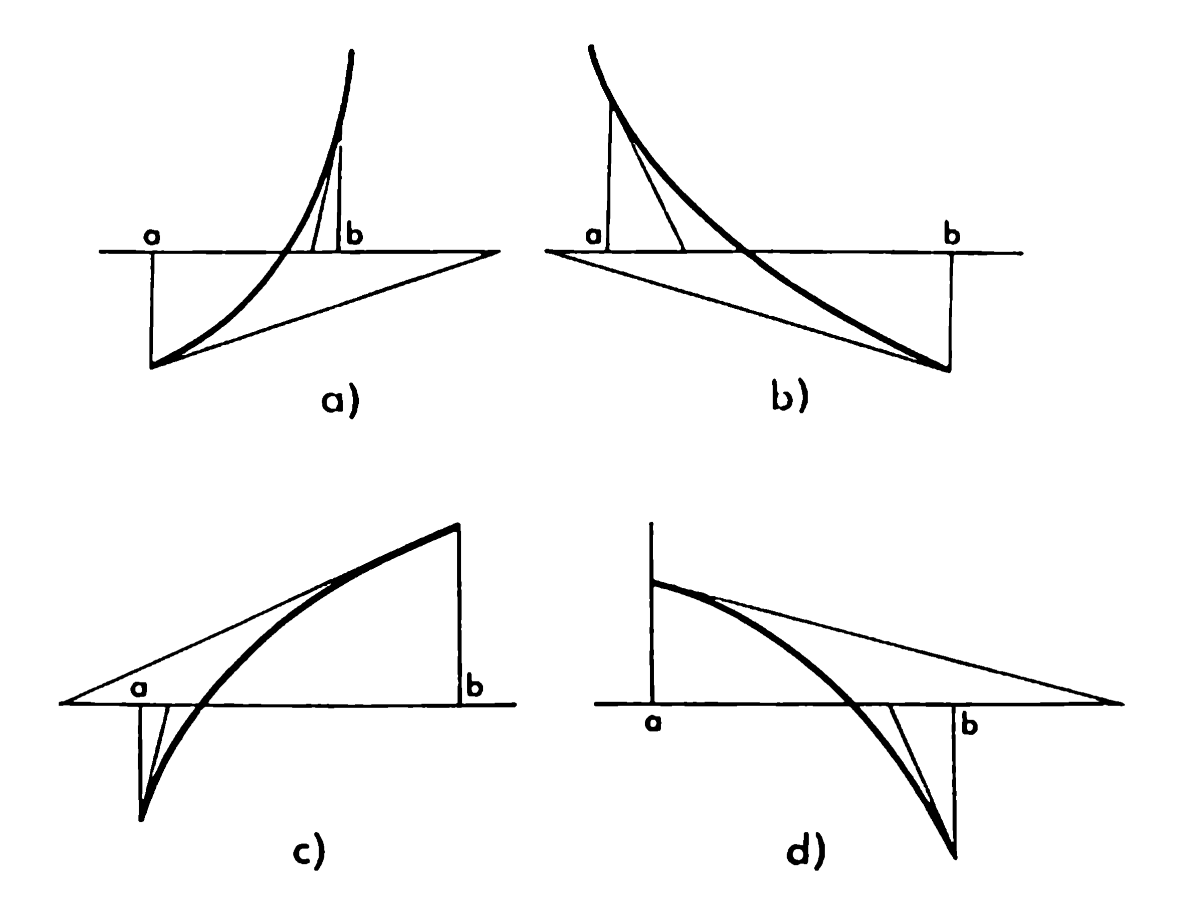


Fig. 17

sa dérivée première est

$$f'(x) = 3x^2 - 8x + 3,$$

et sa dérivée seconde

$$f''(x) = 6x - 8.$$

En mathématiques supérieures, on montre que si la dérivée seconde est positive dans l'intervalle $[a, b]$, la courbe est concave vers le bas. Si la dérivée seconde est négative dans l'intervalle $[a, b]$, la courbe est convexe vers le haut. En utilisant ce qui vient d'être dit, on obtient la règle suivante pour la méthode de Newton:

Supposons qu'en a et b la fonction f (x) prenne des valeurs de signes contraires et qu'entre a et b elle admette une déri-

vée seconde positive. Pour première approximation il convient alors de prendre celui des points a ou b où la fonction f (x) est positive. Si dans l'intervalle [a, b] la dérivée seconde est négative, il convient de prendre pour première approximation le point où la fonction f (x) est négative.

§ 20 MÉTHODE COMBINÉE DE RÉSOLUTION D'ÉQUATIONS

Lorsqu'on résout une équation, on combine souvent la méthode des parties proportionnelles et celle de Newton. Si la courbe représentative de la fonction $y = f(x)$ est concave vers le haut, on détermine les points a_1 et x_1 en utilisant les formules

$$a_1 = a - f(a)\frac{b-a}{f(b)-f(a)}, \qquad (78)$$

$$x_1 = b - \frac{f(b)}{f'(b)}. \qquad (79)$$

Si la courbe représentative de la fonction $y = f(x)$ est concave vers le bas, on détermine le point a_1 au moyen de la formule (78) et le point x_1 au moyen de la formule

$$x_1 = a - \frac{f(a)}{f'(a)}. \qquad (80)$$

Comme on le voit sur les fig. 18, *a* et 18, *b*, la racine ξ de l'équation $f(x) = 0$ est habituellement comprise entre les points a_1 et x_1 ainsi obtenus. En appliquant de nouveau à ces points les formules établies dans la méthode des parties proportionnelles et dans celle de Newton, on obtient un nouveau couple de points a_2 et x_2 et ainsi de suite.

En poursuivant cette procédure, on détermine deux suites de points $a_1, a_2, \ldots, a_n, \ldots$ et $x_1, x_2, \ldots, x_n, \ldots$ qui approchent des deux côtés la racine cherchée ξ. L'avantage de cette méthode est qu'elle permet d'obtenir des valeurs approchées aussi bien par défaut que par excès.

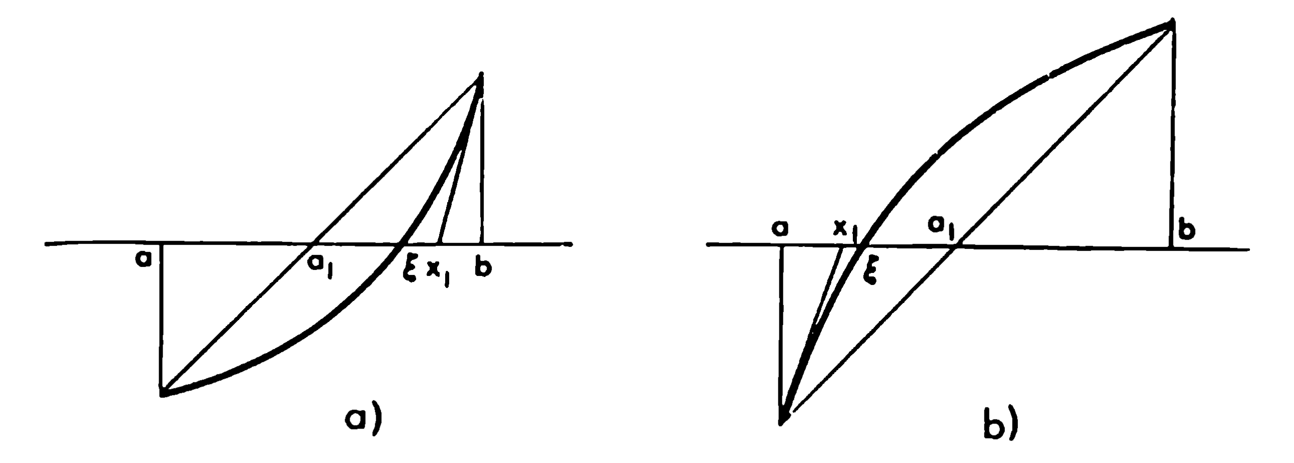


Fig. 18

E x e m p l e. Résoudre avec une précision de 0,001 l'équation

$$x - \sin x - 0{,}5 = 0$$

en utilisant la méthode combinée.

Dressons le tableau des valeurs de la fonction continue

$$f(x) = x - \sin x - 0{,}5.$$

x	-1	0	1	2
$f(x)$	$-0{,}659$	$-0{,}5$	$-0{,}341$	$0{,}591$

On voit sur ce tableau que la solution de l'équation est comprise entre 1 et 2. D'après les formules 2 et 4 du § 18 il vient

$$f'(x) = 1 - \cos x.$$

La formule de Newton s'écrit donc

$$x_{n+1} = x_n - \frac{x_n - \sin x_n - 0{,}5}{1 - \cos x_n}. \qquad (81)$$

Pour savoir laquelle des valeurs 1 ou 2 il convient de prendre pour x_0, calculons la dérivée seconde de la fonction $f(x)$.

D'après la formule 5 § 18, on a $f''(x) = \sin x$. Or, la fonction $\sin x$ est positive [1]) dans l'intervalle [1, 2]. Si donc on applique la règle indiquée plus haut, il faut prendre pour première approximation la valeur 2, car $f(2)$ est positive.

On a donc d'après la formule (81)

$$x_1 = 2 - \frac{2 - \sin 2 - 0{,}5}{1 - \cos 2} = 2 - \frac{2 - 0{,}909 - 0{,}5}{1 + 0{,}416} = 1{,}583.$$

D'autre part, d'après la formule (78) il vient

$$a_1 = 1 - (-0{,}341)\frac{2-1}{0{,}591 - (-0{,}341)} = 1{,}366.$$

En appliquant les formules (81) et (78) à l'intervalle $[a_1, x_1]$ on trouve

$$x_2 = 1{,}583 - \frac{1{,}583 - 1{,}000 - 0{,}5}{1 + 0{,}012} = 1{,}501$$

et

$$a_2 = 1{,}366 + 0{,}113\frac{1{,}583 - 1{,}366}{0{,}083 + 0{,}113} = 1{,}491.$$

En répétant il vient:

$$x_3 = 1{,}498,$$
$$a_3 = 1{,}498.$$

Donc la racine de notre équation est égale à 1,498 à 0,001 près.

§ 21 INDICE DE CONVERGENCE DU PROCESSUS ITÉRATIF

Utilisons maintenant la notion de dérivée pour déduire un nouvel indice de convergence du processus itératif. Nous aurons besoin pour cela d'une formule dite formule de Lagrange (mathématicien français du XVIII[e]).

---

[1]) Sin x est positif dans l'intervalle $[0; \pi] = [0; 3{,}141\ldots]$. Il est donc positif dans la partie [1, 2] de cet intervalle.

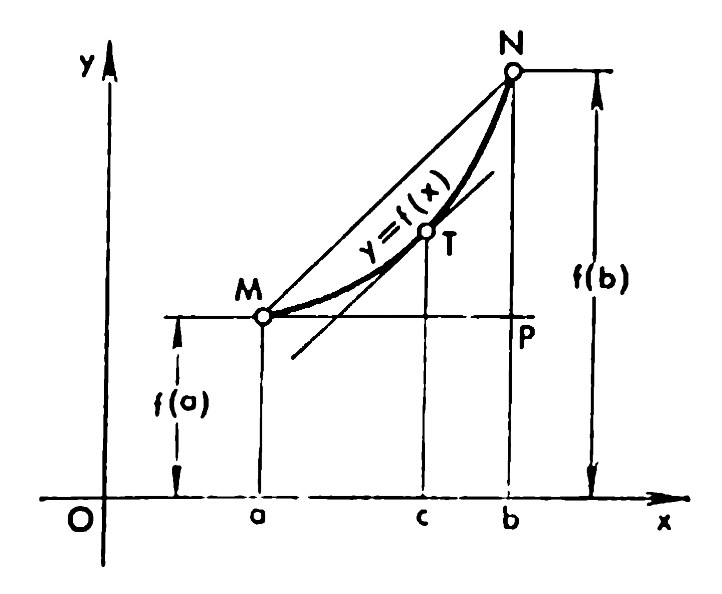


Fig. 19

Considérons la courbe $y = f(x)$ dans l'intervalle $[a, b]$. Désignons par M et N les extrémités de cette courbe et traçons la corde MN (fig. 19). Le coefficient angulaire de cette corde est

$$k_{\text{corde}} = \operatorname{tg} \psi = \frac{PN}{MP}.$$

Or, $MP = b - a$ et $PN = f(b) - f(a)$ donc

$$k_{\text{corde}} = \frac{f(b) - f(a)}{b - a}.$$

Désignons maintenant par T le point de l'arc MN le plus éloigné de la corde MN. Si par ce point on mène une parallèle à la corde, elle sera tangente à la courbe : si cette parallèle coupait la courbe en un point quelconque, celui-ci serait plus éloigné que T. En d'autres termes, la tangente à la courbe en T est parallèle à la corde MN et possède donc le même coefficient angulaire. Or celui-ci est égal à $f'(c)$, où c est l'abscisse du point T. On a donc la formule suivante :

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \qquad (82)$$

Cette formule s'appelle formule de Lagrange. Remarquons que dans cette formule le point c est toujours compris

entre les points a et b. On peut encore écrire cette formule:

$$f(b) - f(a) = f'(c)(b - a). \qquad (83)$$

Revenons maintenant à la résolution de l'équtaion $x = \varphi(x)$ par la méthode des itérations. Supposons que l'application $y = \varphi(x)$ transforme l'intervalle $[a, b]$ en lui-même et qu'elle soit telle que $|\varphi'(x)| < q$, où q est un nombre inférieur à l'unité, $q < 1$. Considérons dans l'intervalle $[a, b]$ deux points quelconques x_1 et x_2. Les points $\varphi(x_1)$ et $\varphi(x_2)$ sont également compris dans l'intervalle $[a, b]$. La formule de Lagrange donne

$$\varphi(x_2) - \varphi(x_1) = \varphi'(c)(x_2 - x_1),$$

où c est un point quelconque compris entre x_1 et x_2 et donc entre a et b. D'autre part, l'inégalité $|\varphi'(c)| < q < 1$, entraîne

$$|\varphi(x_2) - \varphi(x_1)| \leqslant q\,|x_2 - x_1|. \qquad (84)$$

L'inégalité (84) montre que $\varphi(x)$ définit une application contractante. Or nous savons que si $x \to \varphi(x)$ est une application contractante transformant l'intervalle $[a, b]$ en lui-même, quel que soit le point x_0 de cet intervalle, la suite $x_0, x_1, \ldots, x_n, \ldots$ où $x_{n+1} = \varphi(x_n)$ converge vers une solution de l'équation $x = \varphi(x)$. Nous venons donc de démontrer le théorème suivant:

Théorème. *Supposons que la fonction* $y = \varphi(x)$ *définisse une application de l'intervalle* $[a, b]$ *dans lui-même, et que dans cet intervalle soit vérifiée l'inégalité* $|\varphi'(x)| < q$ *où* $q < 1$. *Alors quel que soit le point* x_0 *de l'intervalle* $[a, b]$, *la suite des points* $x_0, x_1, \ldots, x_n, \ldots$, *où* $x_{n+1} = \varphi(x_n)$, *converge vers une racine de l'équation* $x = \varphi(x)$.

Grosso modo, le théorème démontré traduit le fait que le processus itératif permet de trouver ces racines ξ de l'équation $x = \varphi(x)$ pour lesquelles $|\varphi'(\xi)| < 1$. On peut dire que ces points attirent la ligne brisée représentant géométriquement le processus itératif (voir § 8) tandis que les points où $|\varphi'(\xi)| > 1$ la repoussent.

Si l'inégalité $|\varphi'(x)| < q < 1$ est vérifiée sur l'axe des x tout entier, le processus itératif converge quel que soit la première approximation x_0 (voir § 10).

Exemple 1. Peut-on appliquer la méthode des itérations à la résolution de l'équation

$$x = \frac{\cos x + \sin x}{4}\ ?$$

On a

$$\varphi(x) = \frac{\cos x + \sin x}{4},$$

donc

$$\varphi'(x) = \frac{-\sin x + \cos x}{4}.$$

Or, $|\sin x| \leqslant 1$ et $|\cos x| \leqslant 1$, donc

$$|\varphi'(x)| = \left|\frac{-\sin x + \cos x}{4}\right| \leqslant \frac{|\sin x| + |\cos x|}{4} < \frac{1}{2},$$

on peut donc appliquer la méthode des itérations.

Exemple 2. Peut-on appliquer la méthode des itérations à la résolution de l'équation

$$x = 4 - 2^x? \tag{85}$$

La racine cherchée est comprise dans l'intervalle [1, 2], car la fonction continue $y = x - 4 + 2^x$ change de signe dans cet intervalle

$$1 - 4 + 2^1 < 0, \text{ et } 2 - 4 + 2^2 > 0.$$

Nous avons ici

$$\varphi'(x) = -2^x \ln 2.$$

Evaluons l'expression $2^x \ln 2$ dans l'intervalle [1, 2]. Si

$$1 \leqslant x \leqslant 2,$$

alors

$$2 \leqslant 2^x \leqslant 4$$

et donc

$$2 \ln 2 \leqslant 2^x \ln 2 \leqslant 4 \ln 2.$$

La table des logarithmes naturels (dont la base est le nombre $e \approx 2,718 \ldots$) donne pour ln 2 la valeur 0,69 . . . Donc dans l'intervalle [1, 2] a lieu l'inégalité

$$1,38 \ldots \leqslant 2^x \ln 2 \leqslant 2,76 \ldots;$$

on peut donc utiliser la méthode des itérations.

A ces fins mettons l'équation (85) sous la forme

$$2^x = 4 - x.$$

Prenons le logarithme à la base 2 de cette égalité. Nous obtenons

$$x = \log_2 (4 - x).$$

Dans ce cas

$$\varphi'(x) = -\frac{1}{(4-x)\ln 2}$$

et dans l'intervalle [1, 2] est satisfaite l'inégalité

$$|\varphi'(x)| < \frac{1}{2 \ln 2} = \frac{1}{1,38} < 1.$$

(On laisse au lecteur le soin de déduire cette inégalité.)

Le processus itératif est donc convergent.

§ 22 VITESSE DE CONVERGENCE DU PROCESSUS ITERATIF[1])

Prenons maintenant la dérivée de la fonction $\varphi(x)$ pour évaluer la vitesse de convergence du processus itératif dans la résolution de l'équation $x = \varphi(x)$. Nous voulons évaluer la vitesse avec laquelle décroissent les erreurs $\alpha_n = \xi - x_n$ commises sur les valeurs $x_1, \ldots, x_n, \ldots$ de la racine ξ.

[1]) Ce paragraphe peut être omis en première lecture.

Remarquons que sont satisfaites les égalités $\xi = \varphi(\xi)$, $x_{n+1} = \varphi(x_n)$. Il en résulte

$$\alpha_{n+1} = \xi - x_{n+1} = \varphi(\xi) - \varphi(x_n).$$

Or, d'après la formule de Lagrange

$$\varphi(\xi) - \varphi(x_n) = \varphi'(c_n)(\xi - x_n) = \varphi'(c_n)\,\alpha_n,$$

où c_n est un point compris entre les points x_n et ξ. Donc

$$\alpha_{n+1} = \varphi'(c_n)\,\alpha_n. \tag{86}$$

L'égalité (86) entraîne l'assertion suivante:

Supposons que l'équation $x = \varphi(x)$ admette une solution ξ comprise dans l'intervalle $[a, b]$. Si dans cet intervalle est vérifiée l'inégalité $|\varphi'(x)| < q < 1$ et si la première approximation x_1 est aussi comprise dans cet intervalle, alors quel que soit n est satisfaite la relation

$$|\alpha_{n+1}| < q^n |\alpha_1|. \tag{87}$$

En effet, l'égalité (86) entraîne

$$|\alpha_2| = |\varphi'(c_1)|\,|\alpha_1|.$$

Or, le point c_1 est compris dans l'intervalle $[a, b]$ (fig. 20) et donc

$$|\varphi'(c_1)| < q,$$

d'où

$$|\alpha_2| < q\,|\alpha_1|.$$

On établit d'une façon analogue

$$|\alpha_3| = |\varphi'(c_2)|\,|\alpha_2| < q\,|\alpha_2| < q^2\,|\alpha_1|,$$

et en général

$$|\alpha_{n+1}| < q^n |\alpha_1|.$$

Ce qui démontre notre assertion.

Comme pour $0 < q < 1$ la suite des nombres $q, q^2, \ldots, q^n, \ldots$ tend vers 0, l'erreur α_{n+1} tend vers 0 lorsque n croît. En d'autres termes, sous les suppositions faites

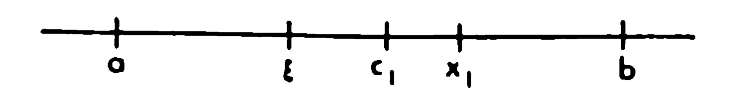


Fig. 20

précédemment les nombres $x_1, \ldots, x_n, \ldots$ se rapprochent du nombre ξ, la différence $|\xi - x_{n+1}|$ décroissant plus vite que $|\alpha_1| q^n$.

De la même façon on démontre que le processus itératif est divergent si l'inégalité

$$|\varphi'(x)| > 1,$$

est satisfaite dans l'intervalle $[a, b]$.

Le processus itératif converge particulièrement vite si la dérivée de la fonction $\varphi(x)$ s'annule en ξ. Dans ce cas, lorsqu'on se rapproche de ξ, la dérivée $\varphi'(x)$ tend vers 0 en vertu de la formule

$$|\alpha_{n+1}| = |\varphi'(c_n)| |\alpha_n|$$

et la convergence du processus s'accélère au fur et à mesure que l'on se rapproche du point ξ.

Ce cas s'est déjà présenté lorsque nous avons appliqué la méthode des itérations à l'extraction de racines carrées. Rappelons que nous avons remplacé l'équation $x^2 = a$ par l'équation $x = \frac{x^2+a}{2x}$. Or la dérivée de la fonction $\varphi(x) = \frac{x^2+a}{2x}$ est :

$$\varphi'(x) = \frac{(x^2+a)' \, 2x - (x^2+a)(2x)'}{4x^2} = \frac{2x \cdot 2x - (x^2+a)\, 2}{4x^2} = \frac{x^2-a}{2x^2}$$

(voir règle 4 § 18 et formule (55) § 13), donc

$$\varphi'(\sqrt{a}) = \frac{(\sqrt{a})^2 - a}{2(\sqrt{a})^2} = 0.$$

Donc la dérivée de la fonction $\varphi(x)$ s'annule au point $x = \sqrt{a}$, ce qui a pour effet d'accélérer la convergence du processus itératif à mesure que l'on se rapproche du point $x = \sqrt{a}$.

Ce phénomène d'accélération du processus itératif à mesure que l'on se rapproche d'une racine de l'équation se rencontre également dans la méthode de Newton (dont la méthode d'extraction de racines est un cas particulier). Nous avons en effet signalé que la méthode de Newton consistait à remplacer l'équation $f(x) = 0$ par l'équation

$$x = x - \frac{f(x)}{f'(x)}$$

que l'on résolvait ensuite par des approximations successives. Nous avons alors

$$\varphi(x) = x - \frac{f(x)}{f'(x)},$$

or

$$\varphi'(x) = 1 - \left[\frac{f(x)}{f'(x)}\right]' = 1 - \frac{f'(x)[f(x)]' - f(x)[f'(x)]'}{[f'(x)]^2} =$$

$$= 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Comme en ξ on a $f(\xi) = 0$, il en résulte que $\varphi'(\xi) = 0$. Or, nous avons vu que cela se traduisait par une accélération de la convergence à l'approche du point ξ.

§ 23 RÉSOLUTION DE SYSTÈMES D'ÉQUATIONS LINÉAIRES PAR LA MÉTHODE DES APPROXIMATIONS SUCCESSIVES

Jusqu'à présent nous n'avons résolu que des équations à une inconnue. Passons à l'étude de systèmes d'équations. Nous commencerons par les systèmes d'équations du premier degré.

Soit m équations du premier degré à m inconnues[1])

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mm}x_m = b_m. \end{array}\right\} \quad (88)$$

Ces systèmes interviennent dans de nombreuses applications. Lorsqu'ils mesurent les dimensions de vastes régions du globe terrestre, les géodésiens sont parfois amenés à résoudre des systèmes de plusieurs centaines d'équations. Il en est de même pour les ingénieurs lorsqu'ils calculent la résistance des constructions métalliques, etc.

La résolution de ces systèmes d'équations par les méthodes habituelles (méthode d'élimination, par exemple) pose souvent d'énormes difficultés. Bien plus commode s'avère la méthode des approximations successives. Illustrons la marche à suivre sur un exemple.

Soit le système d'équations

$$\begin{aligned} 10x_1 - 2x_2 + x_3 &= 9, \\ x_1 + 5x_2 - x_3 &= 8, \\ 4x_1 + 2x_2 + 8x_3 &= 39. \end{aligned}$$

On demande de trouver les inconnues x_1, x_2 et x_3 à 0,01 près.

Explicitons x_1 dans la première équation, x_2 dans la deuxième et x_3 dans la troisième. Le système s'écrit alors:

$$\left.\begin{array}{l} x_1 = 0{,}9 + 0{,}2x_2 - 0{,}1x_3, \\ x_2 = 1{,}6 - 0{,}2x_1 + 0{,}2x_3, \\ x_3 = 4 - 0{,}5x_1 - 0{,}25x_2. \end{array}\right\} \quad (89)$$

---

[1]) Les inconnues sont désignées par les lettres $x_1, x_2, \ldots, x_m$ et les coefficients par les lettres a_{ij}. Le premier indice indique le numéro de l'équation, le second le numéro de l'inconnue. Ainsi, a_{47} représente dans la $4^{\text{ième}}$ équation le coefficient de x_7.

Pour premières approximations prenons des valeurs quelconques de x_1, x_2 et x_3 par exemple $x_1^{(0)} = 0$, $x_2^{(0)} = 0$, $x_3^{(0)} = 0$. Portons ces valeurs dans les seconds membres ; nous obtenons alors les approximations :

$$x_1^{(1)} = 0,9 ;$$
$$x_2^{(1)} = 1,6 ;$$
$$x_3^{(1)} = 4.$$

Portons les valeurs obtenues dans les seconds membres des équations (89). Nous obtenons les approximations

$$x_1^{(2)} = 0,9 + 0,2 \cdot 1,6 - 0,1 \cdot 4 = 0,82,$$
$$x_2^{(2)} = 1,6 - 0,2 \cdot 0,9 + 0,2 \cdot 4 = 2,22,$$
$$x_3^{(2)} = 4 - 0,5 \cdot 0,9 - 0,25 \cdot 1,6 = 3,15.$$

D'une façon générale, sachant les valeurs $x_1^{(n)}$, $x_2^{(n)}$, $x_3^{(n)}$ on trouve les approximations suivantes en appliquant les formules

$$\left.\begin{aligned} x_1^{(n+1)} &= 0,9 + 0,2x_2^{(n)} - 0,1x_3^{(n)}, \\ x_2^{(n+1)} &= 1,6 - 0,2x_1^{(n)} + 0,2x_3^{(n)}, \\ x_3^{(n+1)} &= 4 - 0,5x_1^{(n)} - 0,25x_2^{(n)}. \end{aligned}\right\} \qquad (90)$$

Les résultats des calculs figurent sur le tableau 2.

Tableau 2

n	1	2	3	4	5	6
$x_1^{(n)}$	0,9	0,82	1,03	1,01	1,00	1,00
$x_2^{(n)}$	1,6	2,22	2,07	2,00	1,99	2,00
$x_3^{(n)}$	4,0	3,15	3,03	2,97	3,00	3,00

Nous constatons que dans les limites de la précision demandée sont satisfaites les égalités:

$$x_1^{(5)} = x_1^{(6)}, \quad x_2^{(5)} = x_2^{(6)}, \quad x_3^{(5)} = x_3^{(6)}. \qquad (91)$$

En faisant $n = 5$ dans les égalités (90) et en tenant compte de (91), on obtient dans les limites de la précision demandée:

$$x_1^{(5)} \approx 0{,}9 + 0{,}2x_2^{(5)} - 0{,}1x_3^{(5)},$$
$$x_2^{(5)} \approx 1{,}6 - 0{,}2x_1^{(5)} + 0{,}2x_3^{(5)},$$
$$x_3^{(5)} \approx 4 - 0{,}5x_1^{(5)} - 0{,}25x_2^{(5)}$$

(en réalité, ces égalités se vérifient exactement, mais l'essentiel n'est pas là). Il en résulte que $x_1^{(5)} = 1{,}00$; $x_2^{(5)} = 2{,}00$; $x_3^{(5)} = 3{,}00$ sont (toujours avec une précision de 0,01) solutions du système d'équations donné.

On procède de même dans le cas général [1]). Soit le système d'équations (88). Explicitons x_1 dans la première équation, x_2 dans la deuxième et ainsi de suite. Le système (88) s'écrit alors

$$\begin{aligned}
x_1 &= \frac{b_1}{a_{11}} \qquad\qquad - \frac{a_{12}}{a_{11}} x_2 - \ldots - \frac{a_{1m}}{a_{11}} x_m,\\
x_2 &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1 - \qquad\qquad \ldots - \frac{a_{2m}}{a_{22}} x_m,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
x_m &= \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1 - \frac{a_{m2}}{a_{mm}} x_2 \quad \ldots - \frac{a_{m,\,m-1}}{a_{mm}} x_{m-1}.
\end{aligned} \qquad (92)$$

Soient $x_1^{(1)}, \ldots, x_m^{(1)}$ des premières approximations quelconques des inconnues $x_1, \ldots, x_m$.

En portant ces approximations dans les seconds membres des égalités (92) on obtient les deuxièmes approximations

---

[1]) Toute la suite du paragraphe pourra être omise en première lecture.

$x_1^{(2)}, \ldots, x_m^{(2)}$ des inconnues:

$$x_1^{(2)} = \frac{b_1}{a_{11}} \qquad\qquad - \frac{a_{12}}{a_{11}} x_2^{(1)} - \ldots - \frac{a_{1m}}{a_{11}} x_m^{(1)},$$

$$x_2^{(2)} = \frac{b_1}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(1)} - \qquad\qquad \ldots - \frac{a_{2m}}{a_{22}} x_m^{(1)},$$

. . . .

$$x_m^{(2)} = \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1^{(1)} - \qquad\qquad \ldots - \frac{a_{m,\,m-1}}{a_{mm}} x_{m-1}^{(1)}.$$

D'une façon analogue, sachant les n-ièmes approximations $x_1^{(n)}, x_2^{(n)}, \ldots, x_m^{(n)}$, on obtient les $(n+1)$-ièmes approximations en appliquant les formules:

$$x_1^{(n+1)} = \frac{b_1}{a_{11}} - \qquad\qquad - \frac{a_{12}}{a_{11}} x_2^{(n)} - \ldots - \frac{a_{1m}}{a_{11}} x_m^{(n)},$$

$$x_2^{(n+1)} = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(n)} - \qquad\qquad \ldots - \frac{a_{2m}}{a_{22}} x_m^{(n)},$$

. . . .

$$x_m^{(n+1)} = \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1^{(n)} - \qquad\qquad \ldots - \frac{a_{m,\,m-1}}{a_{mm}} x_{m-1}^{(n)}. \qquad (93)$$

Considérons maintenant le problème inverse c'est-à-dire celui qui consiste, par la résolution d'un système d'équations, à définir l'état final d'un processus itératif.

Soient trois seaux, le premier contenant 12 *litres d'eau, les deux autres étant vides. On verse la moitié du premier seau dans le second, puis la moitié du second dans le troisième et, enfin, la moitié du troisième dans le premier. On répète* 20 *fois ce cycle. On demande de déterminer* (*à* 0,0001 l *près*) *la quantité d'eau contenue dans chaque seau.*

Il est clair qu'il s'agit ici d'approcher un état final de distribution de l'eau. Cette distribution est caractérisée par le fait qu'elle ne varie pas au bout d'un cycle. Expliquons-nous. Si au début d'un cycle le premier seau contient x litres d'eau, le second y et le troisième $12 - x - y$ litres (la quantité totale d'eau ne varie pas lorsqu'on verse l'eau d'un

seau dans un autre), le cycle décrit plus haut est consigné dans le tableau suivant:

	1er seau	2e seau	3e seau
Etat initial	x	y	$12-x-y$
1er transvasement	$x/2$	$x/2+y$	$12-x-y$
2ème transvasement	$x/2$	$x/4+y/2$	$12-3/4x-y/2$
3ème transvasement	$6+\frac{x}{8}-y/4$	$x/4+y/2$	$6-\frac{3}{8}x-y/4$

Ce cycle ne modifiant pas la quantité d'eau contenue dans chaque seau, on a toujours

$$x=6+\frac{x}{8}-\frac{y}{4},$$

$$y=\frac{x}{4}+\frac{y}{2}.$$

On trouve $x = 6$, $y = 3$. Ce qui signifie qu'à la fin des 20 cycles on aura 6 l dans le premier seau, 3 l dans le second et donc 3 l dans le troisième.

Voyons à quelle vitesse une distribution donnée se rapproche de la distribution finale. Supposons que le premier seau contienne a et le second b litres d'eau. Après un cycle on aura dans le premier seau

$$a_1=6+\frac{a}{8}-\frac{b}{4} \tag{94}$$

et dans le second

$$b_1=\frac{a}{4}+\frac{b}{2}. \tag{95}$$

Désignons $a - 6$ par α, $b - 3$ par β, $a_1 - 6$ par α_1 et enfin $b_1 - 3$ par β_1. Il résulte alors des égalités (94) et (95):

$$\alpha_1=a_1-6=\frac{a-6}{8}-\frac{b-3}{4}=\frac{\alpha}{8}-\frac{\beta}{4}$$

et

$$\beta_1=b_1-3=\frac{a-6}{4}+\frac{b-3}{2}=\frac{\alpha}{4}+\frac{\beta}{2}.$$

Après le deuxième cycle, les erreurs sont fournies par les formules:

$$\alpha_2 = \frac{\alpha_1}{8} - \frac{\beta_1}{4} = \frac{1}{8}\left(\frac{\alpha}{8} - \frac{\beta}{4}\right) - \frac{1}{4}\left(\frac{\alpha}{4} + \frac{\beta}{2}\right) = -\frac{3}{64}\alpha - \frac{5}{32}\beta,$$

$$\beta_2 = \frac{\alpha_1}{4} + \frac{\beta_1}{2} = \frac{1}{4}\left(\frac{\alpha}{8} - \frac{\beta}{4}\right) + \frac{1}{2}\left(\frac{\alpha}{4} + \frac{\beta}{2}\right) = \frac{5}{32}\alpha - \frac{3}{16}\beta.$$

Si donc $|\alpha| < \varepsilon$ et $|\beta| < \varepsilon$, alors

$$|\alpha_2| < \frac{13}{64}\varepsilon \simeq 0{,}2\varepsilon,$$

$$|\beta_2| < \frac{11}{32} \simeq 0{,}34\varepsilon.$$

Ce qui signifie qu'après deux cycles, les erreurs α et β diminuent d'au moins trois fois. Donc, après 20 cycles, elles diminuent d'au moins $3^{10} \simeq 70\,000$ fois. Donc après 20 cycles on aura à 0,0001 litre près 6 l dans le premier seau, 3 l dans le second et 3 l dans le troisième.

§ 24 RÉSOLUTION DE SYSTÈMES D'ÉQUATIONS NON LINÉAIRES PAR LA MÉTHODE DES APPROXIMATIONS SUCCESSIVES

La méthode des approximations successives (d'itérations) peut être appliquée avec succès à la résolution de certains systèmes d'équations non linéaires. Soit par exemple le système d'équations:

$$\left.\begin{aligned} x &= 2 + \frac{x^2 + y}{20}, \\ y &= 1 + \frac{x + y^2}{20}. \end{aligned}\right\} \tag{96}$$

Prenons $x_0 = 0$ et $y_0 = 0$ pour premières approximations. Substituons ces valeurs à x et y dans les seconds membres des équations. Nous obtenons $x_1 = 2$, $y_1 = 1$. Portons ces valeurs

dans les seconds membres des équations (96). Nous obtenons

$$x_2 = 2 + \frac{2^2 + 1}{20} = 2,25,$$

$$y_2 = 1 + \frac{2 + 1^2}{20} = 1,15.$$

Poursuivons ce processus:

$$x_3 = 2 + \frac{2,25^2 + 1,15}{20} = 2,31,$$

$$y_3 = 1 + \frac{2,25 + 1,15^2}{20} = 1,18,$$

$$x_4 = 2 + \frac{2,31^2 + 1,18}{20} = 2,33,$$

$$y_4 = 1 + \frac{2,31 + 1,18^2}{20} = 1,18,$$

$$x_5 = 2 + \frac{2,33^2 + 1,18}{20} = 2,33,$$

$$y_5 = 1 + \frac{2,33 + 1,18^2}{20} = 1,18.$$

Nous voyons que les égalités $x_4 = x_5 = 2,33$ et $y_4 = y_5 = 1,18$ sont satisfaites à 0,01 près. On a donc $x = 2,33$ et $y = 1,18$ à 0,01 près.

D'une façon générale, si est donné un système d'équations

$$\left.\begin{aligned} x &= \varphi(x, y) \\ y &= \psi(x, y) \end{aligned}\right\} \qquad (97)$$

où $\varphi(x, y)$ et $\psi(x, y)$ sont des fonctions quelconques, on fait choix d'approximations premières x_0 et y_0 que l'on porte dans les seconds membres des équations (97), puis on applique les formules:

$$\left.\begin{aligned} x_{n+1} &= \varphi(x_n, y_n), \\ y_{n+1} &= \psi(x_n, y_n). \end{aligned}\right\} \qquad (98)$$

Si, pour un indice n, on a $x_{n+1} \approx x_n$, $y_{n+1} \approx x_n$ avec la précision désirée, alors $x \approx x_n$, $y \approx y_n$ avec la même précision.

Les systèmes d'équations à trois inconnues et plus se résolvent d'une façon analogue.

Voyons maintenant sous quelles conditions on est assuré de la convergence des approximations successives dans la résolution des systèmes d'équations.

Nous supposerons que les fonctions $\varphi(x, y)$ et $\psi(x, y)$ intervenant dans le système d'équations (97) sont définies dans un domaine fermé borné D du plan (x, y). En d'autres termes, nous supposerons que le domaine D est entièrement contenu dans un carré et que tous ses points frontières lui appartiennent. Comme exemples de domaines fermés bornés citons le cercle, le polygone (pris avec ses côtés), l'ellipse, etc. Nous supposerons de plus que les fonctions $\varphi(x, y)$ et $\psi(x, y)$ définissent une application du domaine D dans un autre domaine du plan (x, y). Pour trouver le point qui correspond au point $M(x_0, y_0)$, il faut porter ses coordonnées dans $\varphi(x, y)$ et $\psi(x, y)$. Cette substitution nous fournit les coordonnées de l'image M_0. Si par exemple

$$\varphi(x, y) = x^2 + y^2,$$
$$\psi(x, y) = 2xy,$$

l'image du point $M_0(1, 3)$ est le point $N_0(10, 6)$.

Dans la suite nous désignerons l'application définie par les fonctions $\varphi(x, y)$ et $\psi(x, y)$ par la lettre grecque Φ, l'image du point M par $\Phi(M)$, et, enfin, l'image du domaine D par $\Phi(D)$.

Supposons que l'application Φ associe au domaine D sa partie $D_1 = \Phi(D)$. Cette application associe au domaine D_1 sa partie $D_2 = \Phi(D_1)$, qui, de toute évidence, est contenue dans D. En poursuivant cette procédure, on obtient un système de domaines D, D_1, . . ., D_n, . . . (fig. 21) inclus l'un dans l'autre.

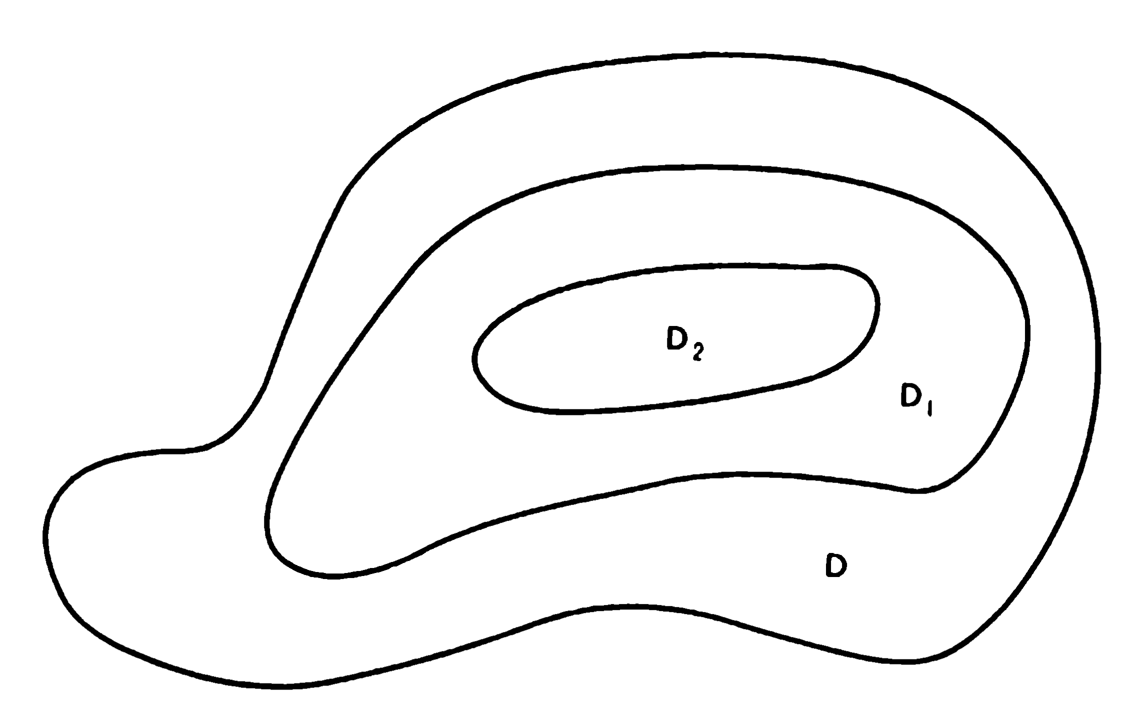


Fig. 21

Nous dirons que l'application Φ est *contractante* par définition s'il existe un nombre q, $0 < q < 1$, tel que quels que soient les points M_1 et M_2 du domaine D on ait la relation :

$$r(\Phi(M_1), \quad \Phi(M_2)) \leqslant qr(M_1, M_2),$$

où $r(M_1, M_2)$ désigne la distance des points M_1 et M_2.

Comme pour le cas d'une variable on démontre la proposition suivante :

Supposons que l'application Φ soit contractante et qu'elle associe au domaine D une de ses parties. Il existe alors dans le domaine D un point N et un seul tel que $N = \Phi(N)$. Ce point appartient à tous les domaines D_n et ses coordonnées ξ et η vérifient le système d'équations (97), *c'est-à-dire*

$$\xi = \varphi(\xi, \eta),$$

$$\eta = \psi(\xi, \eta).$$

Comme dans le cas d'une variable, les valeurs ξ et η sont calculées approximativement par la méthode des itéra-

tions. Si M_0 (x_0, y_0) est un point quelconque du domaine D et $M_{n+1} = \Phi(M_n)$ (c'est-à-dire $x_{n+1} = \varphi(x_n, y_n)$, $y_{n+1} = \psi(x_n, y_n)$), la suite des points $M_0, M_1, \ldots, M_n, \ldots$ converge vers le point fixe N de l'application Φ.

§ 25 DIVERSES DÉFINITIONS DE LA DISTANCE

Le fait que l'application Φ est contractante est un indice suffisant de convergence du processus itératif. Cependant, cet indice n'est pas nécessaire. En effet, l'application Φ peut ne pas être contractante et le processus itératif être convergent. Par exemple l'application Φ définie par les fonctions $\varphi(x, y) = 1 + 2y$, $\psi(x, y) = 3 + \frac{x}{8}$ n'est pas contractante. Si l'on prend les points A (8, 0) et B (8, 4) on a:

$$r(A, B) = 4 \text{ et } \Phi(A) = (1, 4),\ \Phi(B) = (9, 4)$$

et

$$r(\Phi(A), \Phi(B)) = 8 > r(A, B).$$

Cependant, quel que soit le point M_0 considéré, la suite des points $M_0, M_1, \ldots, M_n, \ldots$ converge vers le point

$$N\left(9\frac{1}{3},\ 4\frac{1}{6}\right).$$

On arrive quelquefois à établir la convergence du processus itératif en modifiant la définition de la distance de deux points du plan. En effet, la distance de deux points peut être définie de plusieurs façons. Le voyageur mesurera une distance AB par le temps qu'il mettra pour se rendre de A en B. Lorsque la distance des points A et B (voir fig. 22, *a*) est la somme des segments AC, CD et DB (pour se rendre de A en B il faut atteindre le pont CD, le traverser, puis aller jusqu'en B). Si, sur le plan, on ne peut se déplacer que suivant deux directions perpendiculaires, il y a lieu de prendre pour distance entre les points A et B (voir fig. 22, *b*) la somme des segments AC et CB. Dans d'autres cas il est

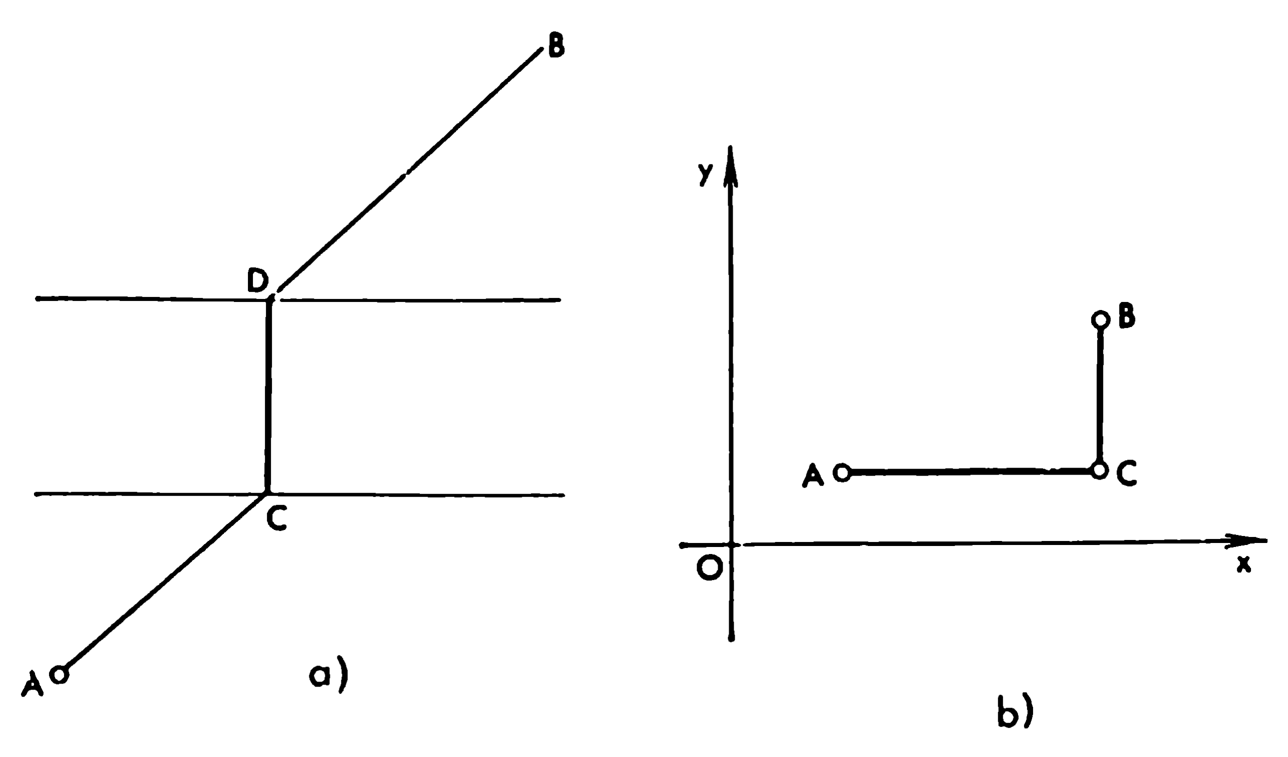


Fig. 22

commode de prendre pour « distance » des points A et B la longueur du plus grand des segments AC et CB. On pourrait imaginer d'autres définitions de la « distance » de deux points. (Pour plus de détails le lecteur pourra se reporter à l'ouvrage de Y. Schreider « Qu'est-ce que la distance ?»).

D'habitude on implique à la distance $r(A, B)$ de jouir des propriétés suivantes:

1) La distance $r(A, B)$ de deux points quelconques A et B *n'est pas négative*; elle n'est égale à zéro que lorsque les points A et B sont confondus.

2) Quels que soient les points A et B, est vérifiée la *condition de symétrie*

$$r(A, B) = r(B, A).$$

3. Quels que soient les points A, B et C, on a l'*inégalité du triangle*

$$r(A, B) \leqslant r(A, C) + r(C, B).$$

Si, dans un ensemble quelconque d'objets, est définie une distance douée de ces propriétés, on dit que cet ensemble est un *espace métrique* et ses éléments, les points de cet espace. Il existe des espaces dont les points sont des fonctions. La distance de deux fonctions $\varphi(x)$ et $\psi(y)$ continues dans l'intervalle $[a, b]$ est définie comme le maximum de la fonction $|\varphi(x) - \psi(x)|$ dans cet intervalle:

$$r(\varphi, \psi) = \max_{a \leqslant x \leqslant b} |\varphi(x) - \psi(x)|.$$

Nous avons vu plus haut qu'un plan peut être muni d'une structure d'espace métrique de plusieurs manières: les propriétés 1)—3) sont valables pour toutes les définitions qu'on a données de la distance.

Indiquons un exemple intéressant où les conditions 1) et 3) sont satisfaites, alors que la condition de symétrie 2) ne l'est pas. Si l'on évalue la distance de deux villes A et B situées au bord d'un fleuve par le temps mis pour parcourir AB, on aura $r(A, B) \neq r(B, A)$ selon que l'on descend ou remonte ce fleuve.

Pour que la méthode des approximations successives soit convergente lorsqu'on résout le système d'équations

$$\left.\begin{aligned} x &= \varphi(x, y), \\ y &= \psi(x, y) \end{aligned}\right\} \tag{99}$$

il suffit que soit remplie la condition suivante:

L'application Φ définie par les fonctions $\varphi(x, y)$ et $\psi(x, y)$ et transformant le domaine D en lui-même doit être contractante relativement à une « distance » $r(A, B)$ au moins. Autrement dit, il doit exister un nombre q, $0 < q < 1$, tel que quels que soient les points M_1 et M_2 de D soit satisfaite l'inégalité

$$r(\Phi(M_1), \Phi(M_2)) \leqslant q r(M_1, M_2).$$

Soit par exemple les fonctions $\varphi(x, y) = 1 + 2y$ et $\psi(x, y) = 3 + \frac{x}{8}$. L'application Φ que ces fonctions définissent est contractante dans le rapport 1/2 si la distance

$r(A, B)$ entre les points $A(x_1, y_1)$ et $B(x_2, y_2)$ est fournie par la formule

$$r(A, B) = \left| \frac{1}{2}(x_2 - x_1) + 2(y_2 - y_1) \right| + \left| \frac{1}{2}(x_2 - x_1) - 2(y_2 - y_1) \right|.$$

Pour les points $A(8,0)$ et $B(8,4)$, on a $r(A, B) = 16$. Pour leurs images $\Phi(A)$ et $\Phi(B)$ on a

$$r(\Phi(A), \Phi(B)) = 8.$$

L'application Φ a réduit de deux fois la « distance ». C'est justement ce qui explique que le processus des itérations appliqué à la résolution du système d'équations

$$x = 1 + 2y$$
$$y = 3 + \frac{x}{8}$$

est convergent bien que l'application Φ ne soit pas contractante relativement à la distance dans son acception habituelle (voir p. 100).

§ 26 INDICE DE CONVERGENCE DE LA MÉTHODE DES APPROXIMATIONS SUCCESSIVES DANS LE CAS DE SYSTÈMES D'ÉQUATIONS LINÉAIRES

Appliquons l'indice de convergence établi plus haut à la résolution de systèmes d'équations linéaires. En faisant choix d'une « distance », nous obtenons des indices de convergence qui sont exprimés en fonction des coefficients de ces systèmes.

Considérons tout d'abord un système de deux équations linéaires à deux inconnues

$$\left.\begin{array}{l} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22}y = b_2. \end{array}\right\} \qquad (100)$$

Supposons $a_{11} \neq 0$ et $a_{22} \neq 0$. En explicitant x dans la première équation et y dans la seconde, on obtient le système suivant

$$x = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} y,$$

$$y = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x.$$

Pour abréger, posons

$$\frac{b_1}{a_{11}} = \beta_1, \quad -\frac{a_{12}}{a_{11}} = \alpha_1, \quad \frac{b_2}{a_{12}} = \beta_2, \quad -\frac{a_{21}}{a_{22}} = \alpha_2.$$

Le système prend alors la forme

$$\left.\begin{aligned} x &= \alpha_1 y + \beta_1, \\ y &= \alpha_2 x + \beta_2. \end{aligned}\right\} \qquad (100')$$

Les fonctions définissant l'application Φ s'écrivent dans ce cas:

$$\varphi(x, y) = \alpha_1 y + \beta_1, \qquad \psi(x, y) = \alpha_2 x + \beta_2.$$

Voyons maintenant à quelles conditions doivent satisfaire les coefficients α_1 et α_2 pour que cette application soit contractante.

On sait que la distance $r(A, B)$ de deux points $A(x_1, y_1)$ et $B(x_2, y_2)$ est exprimée par la formule

$$r(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

L'application Φ associe au point A le point $A_1(\alpha_1 y_1 + \beta_1, \alpha_2 x_1 + \beta_2)$ et au point B le point $B_1(\alpha_1 y_2 + \beta_1, \alpha_2 x_2 + \beta_2)$. La distance de ces deux points est:

$$r(A_1, B_1) = \sqrt{(\alpha_1 y_2 - \alpha_1 y_1)^2 + (\alpha_2 x_2 - \alpha_2 x_1)^2} =$$
$$= \sqrt{\alpha_1^2 (y_2 - y_1)^2 + \alpha_2^2 (x_2 - x_1)^2}. \qquad (101)$$

Posons $q = \max(|\alpha_1|, |\alpha_2|)$.

La formule (101) entraîne alors

$$r(A_1, B_1) \leqslant q \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = q r(A, B).$$

Si $q < 1$, l'application Φ est contractante sur le plan tout entier, et donc la méthode des approximations successives est convergente.

Nous venons de démontrer ainsi que lorsque

$$\max(|\alpha_1|, |\alpha_2|) < 1, \tag{102}$$

le processus itératif converge toujours.

Exprimons l'indice de convergence établi en fonction des coefficients du système d'équations (100). On sait que

$$\alpha_1 = -\frac{a_{12}}{a_{11}}, \quad \alpha_2 = -\frac{a_{21}}{a_{22}}.$$

En portant ces valeurs dans (102) on déduit:

Pour que la méthode des approximations successives converge lorsqu'on l'applique à la résolution du système d'équations linéaires

$$a_{11}x + a_{12}y = b_1,$$
$$a_{21}x + a_{22}y = b_2,$$

il suffit que l'on ait

$$\max\left(\left|\frac{a_{12}}{a_{11}}\right|, \left|\frac{a_{12}}{a_{22}}\right|\right) < 1.$$

Cette condition traduit le fait que les coefficients diagonaux doivent être supérieurs aux coefficients non diagonaux situés sur la même ligne. C'est pourquoi quand on résout un système, par exemple:

$$x - 3y = -11,$$
$$6x + y = 10$$

il faut expliciter y dans la première équation et x dans la seconde. Ce qui donne

$$y = \frac{11}{3} + \frac{1}{3}x,$$
$$x = \frac{5}{3} - \frac{y}{6}.$$

Il est parfois utile de transformer préalablement le système d'équations en y remplaçant les inconnues x et y par des inconnues qui leur sont proportionnelles. Soit par exemple le système d'équations

$$\left.\begin{aligned} 12x + y &= 14, \\ 3x - 2y &= -1. \end{aligned}\right\} \tag{103}$$

On a

$$\max\left(\left|\frac{a_{12}}{a_{11}}\right|, \ \left|\frac{a_{21}}{a_{22}}\right|\right) = \max\left(\frac{1}{12}, \ \frac{3}{2}\right) = \frac{3}{2},$$

on voit donc que la condition suffisante de convergence de la méthode n'est pas remplie. Par contre, si l'on pose $x = 1/3\ z$, on obtient le système

$$\begin{aligned} 4z + y &= 14, \\ z - 2y &= -1, \end{aligned} \tag{104}$$

pour lequel

$$\max\left(\left|\frac{a_{12}}{a_{11}}\right|, \ \left|\frac{a_{21}}{a_{22}}\right|\right) = \max\left(\frac{1}{4}, \ \frac{1}{2}\right) = \frac{1}{2}.$$

On peut donc résoudre le système (104) par la méthode des approximations successives.

Certes, il ne viendra à personne l'idée de résoudre des systèmes aussi élémentaires que le système (104) par la méthode des approximations successives. Cependant celle-ci s'avère particulièrement commode lorsqu'il s'agit de résoudre des systèmes faisant intervenir un grand nombre d'inconnues. Les conditions suffisantes de convergence sont à peu de chose près les mêmes pour le système

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}\right\} \tag{105}$$

que pour un système d'équations à deux inconnues. Cependant, les résultats diffèrent selon la définition qui a été donnée de la distance de deux points. On peut notamment affirmer ce qui suit:

La méthode des approximations successives converge lorsqu'on l'applique au système (105) *si est satisfaite au moins l'une des conditions suivantes*:

$$1)\quad \max_i \sum_{j=1}^{n}{}' \left|\frac{a_{ij}}{a_{ii}}\right| < 1, \tag{106}$$

$$2)\quad \max_j \sum_{i=1}^{n}{}' \left|\frac{a_{ij}}{a_{ii}}\right| < 1, \tag{107}$$

$$3)\quad \max_k \sum_{i,\,j=1}^{n}{}^{(k)} \left|\frac{a_{ij}}{a_{ii}}\right|^2 < 1. \tag{108}$$

Dans les sommes (106) et (107) l'accent signifie qu'on omet le terme pour lequel $i = j$. Dans la somme (108) l'indice (k) signifie qu'on omet les termes pour lesquels $i = j$ et $i = k$. Observons que la condition 3) est a priori satisfaite si

$$\sum_{i,\,j=1}^{n}{}' \left|\frac{a_{ij}}{a_{ii}}\right|^2 < 1, \tag{109}$$

où l'accent garde sa signification précédente. Dans la plupart des ouvrages de calcul numérique, cette condition est donnée sous la forme (109).

Nous avons déjà vu que les conditions (106)-(108) sont satisfaites si l'on exige que l'application

$$\begin{array}{l} x_1 \to \dfrac{b_1}{a_{11}} - \dfrac{a_{12}}{a_{11}} x_2 - \ldots - \dfrac{a_{1n}}{a_{11}} x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n \to \dfrac{b_n}{a_{nn}} - \dfrac{a_{n1}}{a_{nn}} x_1 - \dfrac{a_{n2}}{a_{nn}} x_2 - \ldots - \dfrac{a_{1,\,n-1}}{a_{nn}} x_{n-1} \end{array}$$

soit contractante relativement à une distance. Plus exactement, la condition (106) correspond à la distance

$$r(A, B) = \max(|x_1 - y_1|, \ldots, |x_n - y_n|),$$

la condition (107), à la distance $r(A, B) = \sum_{i=1}^{n} |x_i - y_i|$ et la condition (108), à la distance $r(A, B) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$ des points $A(x_1, \ldots, x_n)$ et $B(y_1, \ldots, y_n)$.

Comme dans le cas de deux variables, il est parfois utile de remplacer les inconnues $x_1, \ldots, x_n$ par des quantités proportionnelles $y_1 = p_1 x_1, \ldots, y_n = p_n x_n$, où $p_1 > 0, \ldots, p_n > 0$.

Dans ce cas, les conditions (106), (107) et (108) s'écrivent sous la forme

$$1') \quad \max_i \sum_{j=1}^{n}{}' \left|\frac{a_{ij}}{a_{ii}}\right| \frac{p_i}{p_j} < 1, \qquad (106')$$

$$2') \quad \max_j \sum_{i=1}^{n}{}' \left|\frac{a_{ij}}{a_{ii}}\right| \frac{p_i}{p_j} < 1, \qquad (107')$$

$$3') \quad \max_k \sum_{i,\, j=1}^{n}{}^{(k)} \left|\frac{a_{ij}}{a_{ii}}\right|^2 \frac{p_i^2}{p_j^2} < 1. \qquad (108')$$

En particulier, pour $p_i = |a_{ii}|$ ces coefficients prennent la forme

$$1'') \quad \max_i \sum_{j=1}^{n}{}' \left|\frac{a_{ij}}{a_{jj}}\right| < 1, \qquad (106'')$$

$$2'') \quad \max_j \sum_{i=1}^{n}{}' \left|\frac{a_{ij}}{a_{jj}}\right| < 1, \qquad (107'')$$

$$3'') \quad \max_k \sum_{i,\, j=1}^{n}{}^{(k)} \left|\frac{a_{ij}}{a_{jj}}\right|^2 < 1. \qquad (108'')$$

Illustrons ce que nous venons de développer sur un exemple. Soit à résoudre le système

$$x - 0{,}6y - 0{,}5z = -2{,}6,$$
$$-0{,}2x + y - 0{,}4z = 3,$$
$$-0{,}1x + 0{,}5y + z = 3{,}9.$$

Les conditions 1) et 2) tout comme les conditions 1') et 2') ne sont pas remplies pour ce système. De même n'a pas lieu l'inégalité (109), car la somme des carrés des éléments diagonaux est égale à 1,07. Or,

$$\max_h \sum^{(h)} \left|\frac{a_{ij}}{a_{ii}}\right|^2 = \max(0{,}6^2 + 0{,}5^2 + 0{,}2^2 + 0{,}4^2;\ 0{,}6^2 + 0{,}5^2 + 0{,}1^2 + 0{,}5^2;\ 0{,}2^2 + 0{,}4^2 + 0{,}1^2 + 0{,}5^2) = 0{,}87 < 1,$$

donc on peut appliquer la méthode des approximations successives à la résolution de ce système.

Remarquons que les conditions de convergence de la méthode des approximations successives formulées plus haut sont suffisantes, mais en aucune façon nécessaires. Si l'on choisit une autre définition de la distance de deux points et que l'on écrive la condition de contraction, on obtient d'autres conditions de convergence. Nous n'aborderons pas ce problème ici.

Au sujet des systèmes d'équations linéaires on peut faire les mêmes remarques qu'au § 5. Par exemple, le résultat final est indépendant du choix de la première approximation. Si donc on commet une erreur, celle-ci ne fausse pas les calculs mais retarde simplement le résultat final.

Il existe plusieurs méthodes d'approximations successives pour la résolution de systèmes d'équations linéaires.

Dans certaines d'elles, par exemple, une fois qu'on a trouvé la valeur approchée $x_1^{(n+1)}$ on la porte dans le système en même temps que les valeurs $x_1^{(n+1)}$, $x_3^{(n)}$, $x_4^{(n)}$, . . ., $x_m^{(n)}$ pour déterminer $x_2^{(n+1)}$; ensuite on porte les valeurs $x_1^{(n+1)}$, $x_2^{(n+1)}$, $x_4^{(n)}$, . . ., $x_m^{(n)}$ dans le système pour trouver $x_3^{(n+1)}$ et ainsi de suite. La description de toutes les méthodes des approximations successives est longue et pourrait à elle seule constituer la matière d'un ouvrage.

§ 27 LES APPROXIMATIONS SUCCESSIVES EN GÉOMÉTRIE

Nous avons montré comment on pouvait appliquer la méthode des approximations successives à la résolution d'équations et de systèmes d'équations. Cette méthode se connaît d'autres applications en géométrie et notamment dans la mesure de la longueur de la circonférence. On sait que pour calculer la longueur de la circonférence on calcule d'abord le périmètre d'un carré inscrit, puis celui de l'octogone régulier inscrit, du dioctogone, etc. La limite de ces périmètres est égale précisément à la longueur de la circonférence. Chaque périmètre est calculé à l'aide du périmètre précédent de la manière suivante:

Désignons par A_n le côté d'un polygone régulier de 2^n côtés et par P_n son périmètre. Si A_2 est le côté d'un carré, on a $A_2 = R\sqrt{2}$ et $P_2 = 4R\sqrt{2}$. Supposons P_n connu. Il est alors évident que

$$A_n = \frac{P_n}{2^n}.$$

On démontre en géométrie que le côté a_{2n} d'un polygone régulier inscrit de $2n$ côtés s'exprime en fonction du côté a_n d'un polygone régulier inscrit de n cotés et du rayon R du

cercle[1]),

$$a_{2n} = R\sqrt{2-\sqrt{4-\frac{a_n^2}{R^2}}}. \qquad (110)$$

Donc le côté A_{n+1} d'un polygone régulier inscrit de $2^{(n+1)}$ côtés s'exprime en fonction du côté A_n d'un polygone de 2^n côtés par la formule

$$A_{n+1} = R\sqrt{2-\sqrt{4-\frac{A_n^2}{R^2}}}.$$

Comme $A_n = \frac{P_n}{2^n}$ et $A_{n+1} = \frac{P_{n+1}}{2^{n+1}}$, on a

$$P_{n+1} = 2^{n+1}R\sqrt{2-\sqrt{4-\frac{P_n^2}{2^{2n}R^2}}}. \qquad (110')$$

La suite des nombres $P_2, P_3, \ldots, P_n, \ldots$ tend vers la longueur de la circonférence, c'est-à-dire vers $2\pi R$. On peut

---

[1]) On établit plus aisément cette formule en trigonométrie. Il est évident que si a_n est un côté d'un polygone régulier inscrit de n côtés et a_{2n} celui d'un polygone régulier inscrit de $2n$ côtés, on a

$$a_n = 2R\sin\frac{\pi}{n} \quad \text{et} \quad a_{2n} = 2R\sin\frac{\pi}{2n}$$

(voir fig.) Comme

$$\sin\frac{\alpha}{2} = \sqrt{\frac{1-\cos\alpha}{2}},$$

on a

$$a_{2n} = 2R\sin\frac{\pi}{2n} = 2R\sqrt{\frac{1-\cos\frac{\pi}{n}}{2}} =$$

$$= R\sqrt{2-2\sqrt{1-\sin^2\frac{\pi}{n}}} = R\sqrt{2-\sqrt{4-\frac{a_n^2}{R^2}}}.$$

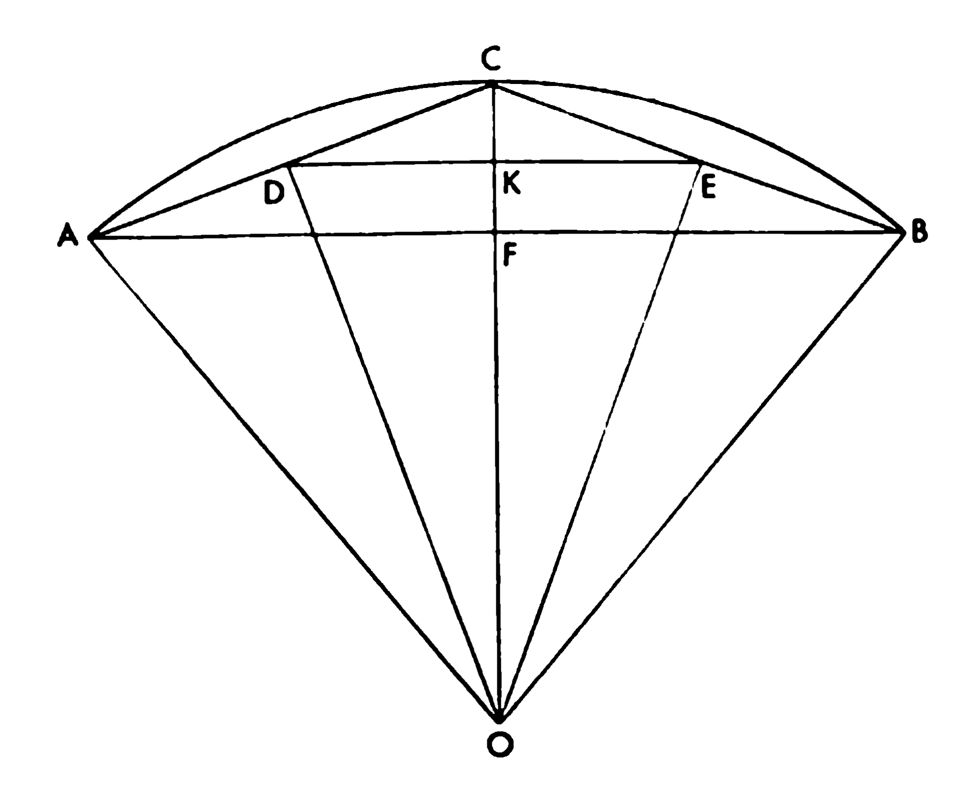


Fig. 23

donc considérer que la formule (110) permet de calculer la valeur de $2\pi R$ par la méthode des applications successives. On peut utiliser cette méthode pour calculer la valeur de π avec autant de décimales que l'on veut.

Il existe une autre méthode approchée pour le calcul de π, appelée *méthode des périmètres égaux*. Dans cette méthode on remplace le polygone régulier de 2^n côtés par un polygone régulier de 2^{n+1} côtés de même périmètre. Désignons par l_n l'apothème du polygone régulier de 2^n côtés et par r_n le rayon de son cercle circonscrit. Désignons d'autre part par l_{n+1} l'apothème du polygone régulier de 2^{n+1} côtés ayant le même périmètre que le polygone de 2^n côtés et par r_{n+1} le rayon de son cercle circonscrit.

Soit AB (fig. 23) un côté du polygone de 2^n côtés inscrit dans le cercle de rayon r_n. Joignons le milieu C de l'arc AB aux points A et B et traçons la ligne médiane DE du triangle ACB. Il est évident que l'angle DOE est égal à la moitié de l'angle AOB. Donc DE est le côté d'un polygone régulier

de 2^{n+1} côtés inscrit dans le cercle de rayon OD. Comme $DE = 1/2\ AB$, le périmètre du polygone de 2^{n+1} côtés est égal à celui du polygone de 2^n côtés. Ce qui signifie que $r_{n+1} = OD$ et $l_{n+1} = OK$.

On établit aisément que

$$l_{n+1} = OK = \frac{r_n + l_n}{2}. \qquad (111)$$

D'autre part, du triangle rectangle ODC on tire la relation

$$r_{n+1} = \sqrt{r_n l_{n+1}}. \qquad (112)$$

Les formules (111) et (112) expriment r_{n+1} et l_{n+1} en fonction de r_n et l_n.

Lorsque n croît, les périmètres des polygones ne varient pas, quant aux nombres r_n et l_n, ils tendent vers la même limite qui est précisément le rayon du cercle dont la longueur est égale au périmètre des polygones considérés. Si on choisit le polygone initial de telle sorte que son périmètre soit égal à 2, r_n et l_n tendent tous deux vers $1/\pi$:

$$\lim_{n\to\infty} r_n = \frac{1}{\pi}; \quad \lim_{n\to\infty} l_n = \frac{1}{\pi}.$$

Si, par exemple, on prend pour premier polygone un carré de côté 1/2, alors $r_2 = \frac{\sqrt{2}}{4}$ et $l_2 = \frac{1}{4}$. On peut donc affirmer ce qui suit: si on pose $r_2 = \frac{\sqrt{2}}{4}$, $l_2 = \frac{1}{4}$ et qu'on calcule les valeurs de r_{n+1}, l_{n+1}, $n = 2, 3, \ldots$ par les formules (111) et (112), alors

$$\lim_{n\to\infty} r_n = \lim_{n\to\infty} l_n = 1/\pi.$$

A l'aide de ces formules on peut trouver approximativement la valeur de $1/\pi$. Pour cela il faudra pousser les calculs jusqu'à ce que les valeurs de r_n et l_n coïncident dans les limites de la précision désirée. Cette valeur commune de r_n et l_n sera prise pour valeur de $1/\pi$ avec la précision demandée.

§ 28 CONCLUSION

Dans cet ouvrage nous avons vu comment appliquer la méthode des approximations successives à la résolution de nombreux problèmes: établissement d'un plan, extraction de racines carrées, résolution d'équations, calcul de la longueur de la circonférence. Cette liste est loin d'être exhaustive. En effet, un très grand nombre de problèmes conduisent aux équations différentielles (qui sont des équations dans lesquelles interviennent les fonctions cherchées et leurs dérivées), aux équations intégrales, ainsi qu'à des équations d'une forme plus compliquée. Et là encore la méthode des approximations successives s'avère la meilleure méthode de résolution approchée. Certes, son utilisation pose ici de plus grandes difficultés que dans la résolution des équations algébriques. Mais, sans crainte d'exagérer, on peut dire que sans la méthode des approximations successives bien des problèmes grandioses auxquels sont confrontées actuellement la physique et plus généralement la technique n'auraient trouvé leurs solutions. Cette méthode est l'âme des calculs de la trajectoire d'un satellite, du fonctionnement d'une pile atomique, de la structure de l'atome, etc. Malheureusement, il nous est impossible d'en brosser un tableau tant soit peu succinct, car il faut faire appel à des notions mathématiques qui débordent le cadre de cet ouvrage.

EXERCICES

Pour que le lecteur puisse vérifier comment il a assimilé les méthodes de résolution approchée des équations nous lui proposons quelques exemples dont il trouvera la solution dans les pages suivantes.

Résoudre par la méthode des itérations les équations[1])

[1]) Dans certains exemples on prendra éventuellement soin de mettre l'équation à résoudre sous la forme $x = \varphi(x)$.

1. $x=\frac{1}{(x+1)^2}$.

2. $x=(x+1)^3$.

3. $x=4+\sqrt[3]{\frac{x-1}{x+1}}$.

4. $x=2\pm\sqrt[4]{x}$.

5. $x=\sqrt[3]{5-x}$.

6. $4-x=\operatorname{tg} x$.

7. $x^2=\sin x$.

8. $x^3=\sin x$.

9. $x=\arcsin\frac{x+1}{4}$.

10. $x=\cos x$.

11. $x=\frac{1}{\cos x}$.

12. $x=1+\frac{1}{10}\sin x$.

13. $x=\pm\sqrt{\lg(x+2)}$.

14. $x^2=\ln(x+1)$.

15. $\ln x=4-x^2$.

16. $\ln x=2-x$.

17. $x^2=e^x+2$.

18. $\lg x=0{,}1x$.

19. $\operatorname{tg} x=\lg x$.

20. $x=\frac{1}{10}e^{-x}$.

Résoudre par la méthode de Newton les équations

21. $x^3-5x+1=0$.

22. $x^3-9x^2+20x-11=0$.

23. $x^3-3x^2-3x+11=0$.

24. $x^5+5x+1=0$.

25. $\sin x+x=1$.

26. $x^2-10\lg x-3=0$.

27. Résoudre à 0,001 près les systèmes d'équations par la méthode des approximations successives:

a) $\begin{cases} x= & 0{,}2y-0{,}1z+0{,}898, \\ y=0{,}3x & +0{,}15z+1{,}383, \\ z=0{,}25x-0{,}4y & +3{,}677; \end{cases}$

b) $\begin{cases} x=\frac{1}{4}\sin(x+y)+0{,}336, \\ y=-\frac{1}{4}\sin(x-y)+0{,}362; \end{cases}$

c) $\begin{cases} x=\sqrt{x+2y}-0{,}710, \\ y=\sqrt{y-x}+1. \end{cases}$

1. Posons $\varphi(x) = \frac{1}{(x+1)^2}$. Alors $\varphi'(x) = \frac{-2}{(1+x)^3}$. On a $\varphi(0) = 1 > 0$, $\varphi(1) = \frac{1}{4} < 1$. La racine de cette équation est donc contenue dans l'intervalle [0, 1]. Cependant, nous ne pourrons pas appliquer la méthode des approximations successives à cet intervalle, car $|\varphi'(0)| = 2 > 1$. Pour réduire cet intervalle, remarquons que $\varphi(0,4) = \frac{1}{1,96} > 0,4$. Donc, cette équation admet une racine dans l'intervalle [0,4; 1]. Si de plus $0,4 \leqslant x \leqslant 1$, alors $|\varphi'(x)| \leqslant \frac{2}{1,4^3} < 1$ et on peut utiliser la méthode des approximations successives. En posant $x_1 = 0,4$ on obtient au bout de 11 pas $x_{11} \approx \varphi(x_{11}) \approx 0,4655$.

On a donc $x = 0,4655$ à 0,0001 près.

2. Posons $\varphi(x) = (x+1)^3$. Alors $\varphi'(x) = 3(x+1)^2$. D'autre part, $\varphi(-2) = -1 > -2$, $\varphi(-3) = -8 < -3$. L'intervalle [−3, −2] contient une racine de cette équation. Nous ne pourrons pas néanmoins appliquer la méthode des approximations successives, car $|\varphi'(x)| > 1$ dans l'intervalle [−3, −2]. Ecrivons l'équation donnée sous la forme

$$x = \sqrt[3]{x} - 1.$$

Alors $\psi(x) = \sqrt[3]{x} - 1$ et $\psi'(x) = \frac{1}{3\sqrt[3]{x^2}}$. Dans l'intervalle [−3, −2] on a $|\psi'(x)| \leqslant \frac{1}{3\sqrt[3]{x^2}} < 1$. On peut donc appliquer la méthode des approximations successives. En posant $x_1 = -2$, il vient $x_6 \approx \approx \psi(x_6) \approx -2,325$. Donc, $x = -2,325$ à 0,001 près.

3. Posons $\varphi(x) = 4 + \sqrt[3]{\frac{x-1}{x+1}}$, d'où

$$\varphi'(x) = \frac{2}{3\sqrt[3]{(x-1)^2(x+1)^4}}.$$

La fig. 24 indique que la droite $y = x$ coupe la courbe $y = 4 + \sqrt[3]{\frac{x-1}{x+1}}$ en deux points contenus respectivement dans les intervalles [−1, 0] et [4, 5]. Dans le premier on a $|\varphi'(x)| \leqslant \frac{2}{15\sqrt[3]{45}} < 1$. En faisant $x_1 = 4$, il vient $x_8 \approx \varphi(x_8) \approx 4,870$. Donc $x_8 = 4,870$ à 0,0001 près.

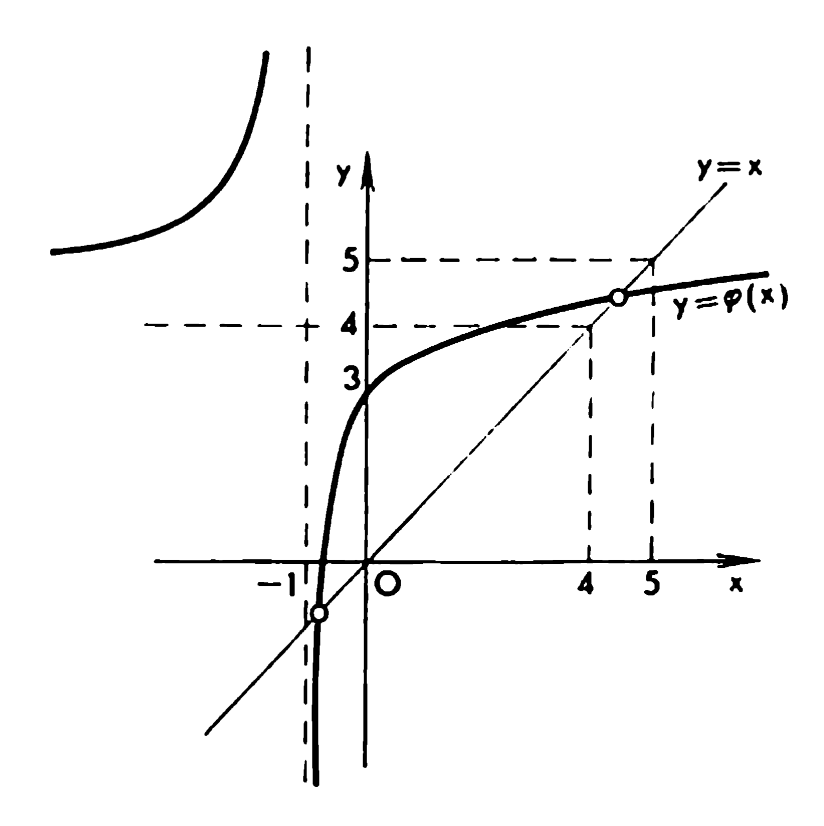


Fig. 24

On ne peut appliquer directement la méthode des approximations successives à l'intervalle $[-1, 0]$. Ecrivons donc l'équation donnée sous la forme $(x-4)^3 = \frac{x-1}{x+1}$, d'où $\frac{x-1}{(x-4)^3} = x+1$ et $x = \frac{x-1}{(x-4)^3} - 1$. Maintenant $\psi(x) = \frac{x-1}{(x-4)^4} - 1$ et $\psi'(x) = \frac{-2x-1}{(x-4)^4}$. Il est clair que pour $-1 \leqslant x \leqslant 0$ on a $|\psi'(x)| \leqslant \frac{1}{256} < 1$, et on peut se servir de la méthode des approximations successives. En posant $x_1 = 0$, il vient $x_2 \approx \psi(x_2) \approx -0{,}9840$. Donc $x = -0{,}9840$ à $0{,}0001$ près. Ceci nous donne deux racines $x = -0{,}9840$ et $x = 4{,}870$.

4. On a $\varphi_1(x) = 2 + \sqrt[4]{x}$ et $\varphi_2(x) = 2 - \sqrt[4]{x}$. Sur la fig. 25, on voit que l'équation $x = 2 + \sqrt[4]{x}$ admet une racine dans l'intervalle $(3, 4)$. Dans cet intervalle

$$|\varphi'(x)| = \frac{1}{4 + \sqrt[4]{x^3}} < 1.$$

Appliquons la méthode des approximations successives. En posant $x_1 = 4$, il vient $x_4 \approx \varphi(x_4) \approx 3{,}353$. Donc $x = 3{,}353$ à $0{,}001$ près.

Résolvons maintenant l'équation $x = 2 - \sqrt[3]{x}$. Elle admet 1 pour racine. Donc les solutions de l'équation sont $x = 1$ et $x = 3{,}353$.

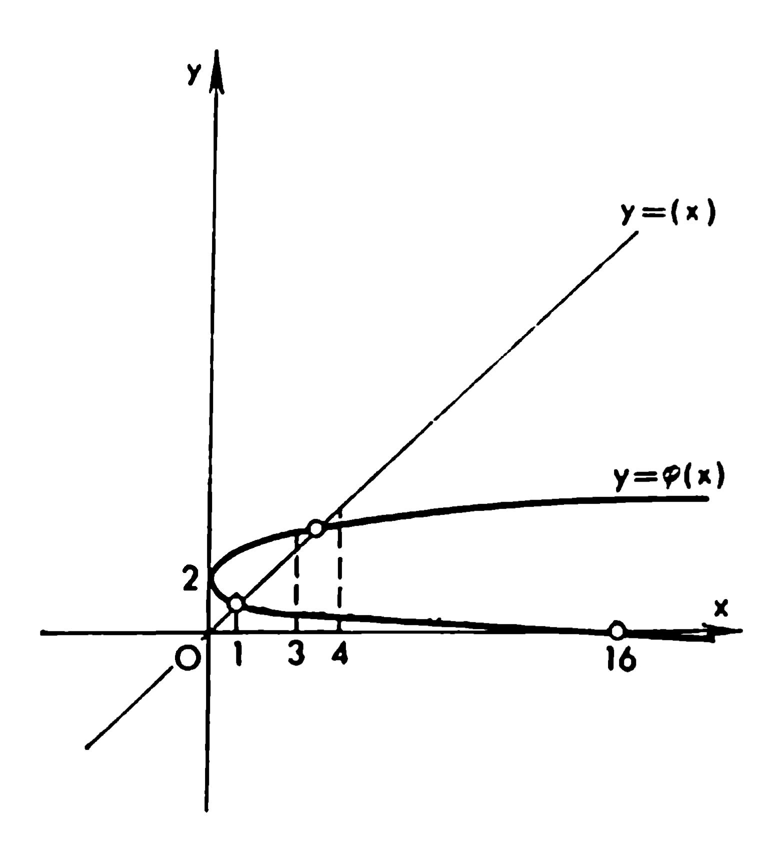


Fig. 25

5. On a $\varphi(x) = \sqrt[3]{5-x}$ et $\varphi'(x) = \dfrac{-1}{3\sqrt[3]{(5-x)^2}}$. D'où

$$\varphi(1) = \sqrt[3]{4} < 2, \quad \varphi(2) = \sqrt[3]{3} > 1.$$

Il existe donc une solution dans l'intervalle [1, 2]. Nous avons dans cet intervalle $|\varphi'(x)| \leqslant \dfrac{1}{3\sqrt[3]{9}} < 1$. En faisant $x_1 = 1$, il vient $x_5 \approx \approx \varphi(x_5) \approx 1{,}516$. Donc, $x = 1{,}516$ à 0,001 près.

6. Ecrivons l'équation sous la forme

$$x = \operatorname{arctg}(4-x).$$

On a $\varphi(x) = \operatorname{arctg}(4-x)$. D'où $\varphi(1) = \operatorname{arctg} 3 \approx 1{,}25$, $\varphi(2) = = \operatorname{arctg} 2 \approx 1{,}10$. Il en résulte qu'il existe une solution de cette équation dans l'intervalle [1, 2]. On a dans cet intervalle

$$|\varphi'(x)| = \frac{1}{1+(4-x^2)} \leqslant \frac{1}{5}.$$

On peut donc appliquer la méthode des approximations successives. En posant $x_1 = 1$, il vient $x_4 \approx \varphi(x_4) \approx 1{,}225$. Donc, on a $x = = 1{,}225$ à 0,001 près.

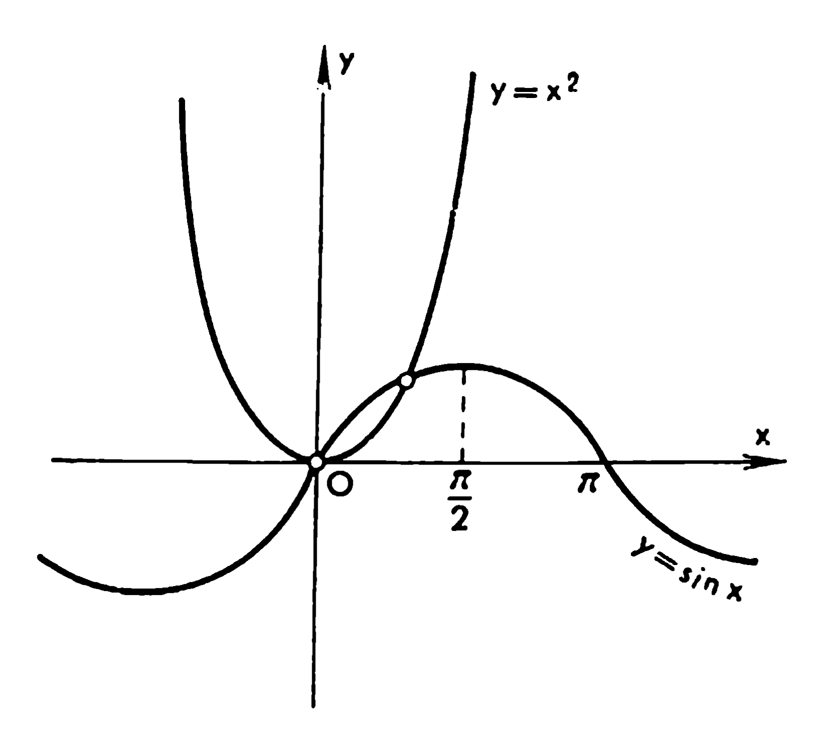


Fig. 26

7. L'équation donnée admet la solution $x = 0$. Sur la fig. 26 on voit que la deuxième solution est positive. Elle vérifie donc l'équation $x = \sqrt{\sin x}$. On a $\varphi(x) = \sqrt{\sin x}$ et $\varphi'(x) = \dfrac{\cos x}{2\sqrt{\sin x}}$. Comme

$$\varphi\left(\frac{1}{2}\right) = \sqrt{\sin {}^1/_2} \approx \sqrt{0{,}4794} < \frac{1}{2}$$

et que

$$\varphi(1) = \sqrt{\sin 1} \approx \sqrt{0{,}8414} < 1,$$

l'équation admet une deuxième solution dans l'intervalle [1/2; 1]. On a dans cet intervalle

$$|\varphi'(x)| \leqslant \frac{\cos \frac{1}{2}}{2\sqrt{\sin {}^1/_2}} \approx \frac{0{,}8703}{1{,}3846} < 1,$$

et la méthode des approximations est convergente. En faisant $x_1 = 1$, il vient $x_7 \approx \varphi(x_7) \approx 0{,}8768$. Donc la deuxième solution est 0,8768 à 0,0001 près.

8. Cet exercice se résout comme le précédent. Ecrivons l'équation sous la forme

$$x = \sqrt[3]{\sin x}.$$

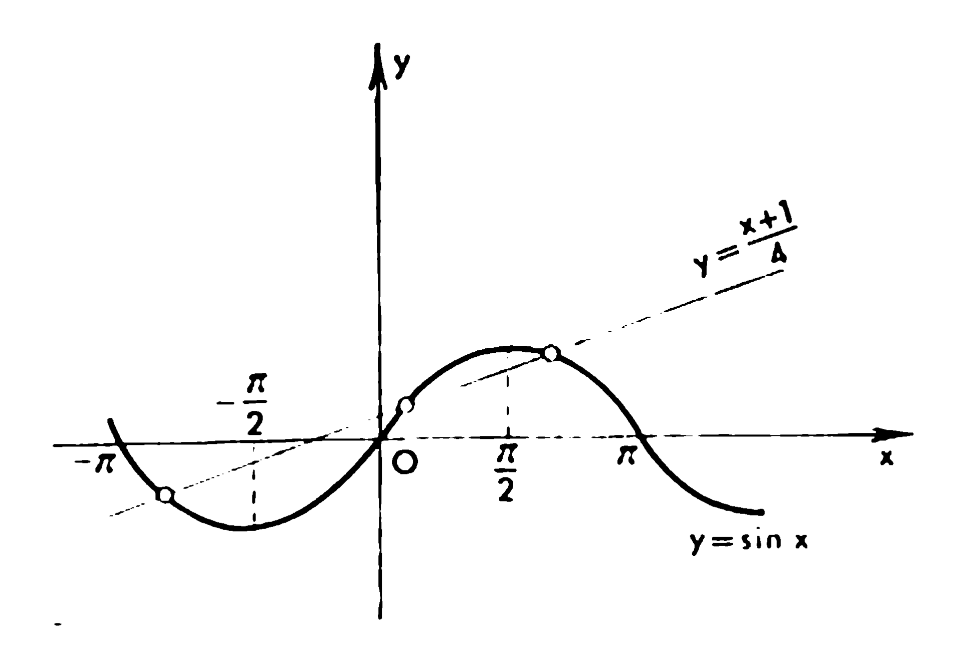


Fig. 27

En prenant $x_1 = 1$ il vient $x_6 \approx \varphi(x_6) \approx 0{,}9286$. On a donc une solution égale à 0,9286 à 0,001 près. Comme les seconds membres de l'équation sont des fonctions impaires, nous avons une deuxième solution égale à $-0{,}9286$. La troisième solution est 0.

9. On peut mettre cette équation sous la forme $\frac{x+1}{4} = \sin x$, $-\pi/2 \leqslant x \leqslant \pi/2$. Sur la fig. 27, on voit que l'équation admet une racine unique comprise entre 0 et $\pi/2$. On a :

$$\varphi(x) = \arcsin \frac{x+1}{4},$$

$$\varphi'(x) = \frac{1}{\sqrt{16-(x+1)^2}}.$$

On a dans l'intervalle $[0, \pi/2]$, $|\varphi'(x)| < \frac{1}{\sqrt{3}} < 1$. En faisant $x=1$ il vient $x_8 \approx \varphi(x_8) \approx 0{,}3422$. Donc $x = 0{,}3422$ à 0,0001 près.

10. Comme $\cos 0 = 1$ et $\cos 1 > 0$, l'équation $x = \cos x$ possède une racine dans l'intervalle $[0, 1]$. Puisque $|\varphi'(x)| \leqslant \sin 1 < < 1$, on peut appliquer la méthode des approximations successives. En faisant $x_1 = 1$ on obtient $x = 0{,}7391$ à 0,0001 près.

11. Sur la fig. 28, on voit que les racines positives de l'équation sont proches des points d'intersection de la fonction $y = \cos x$ avec l'axe des abscisses et sont situées à droite des points d'intersection de la forme $\pi/2 + (2k+1)\pi$ et à gauche des points d'intersection de la

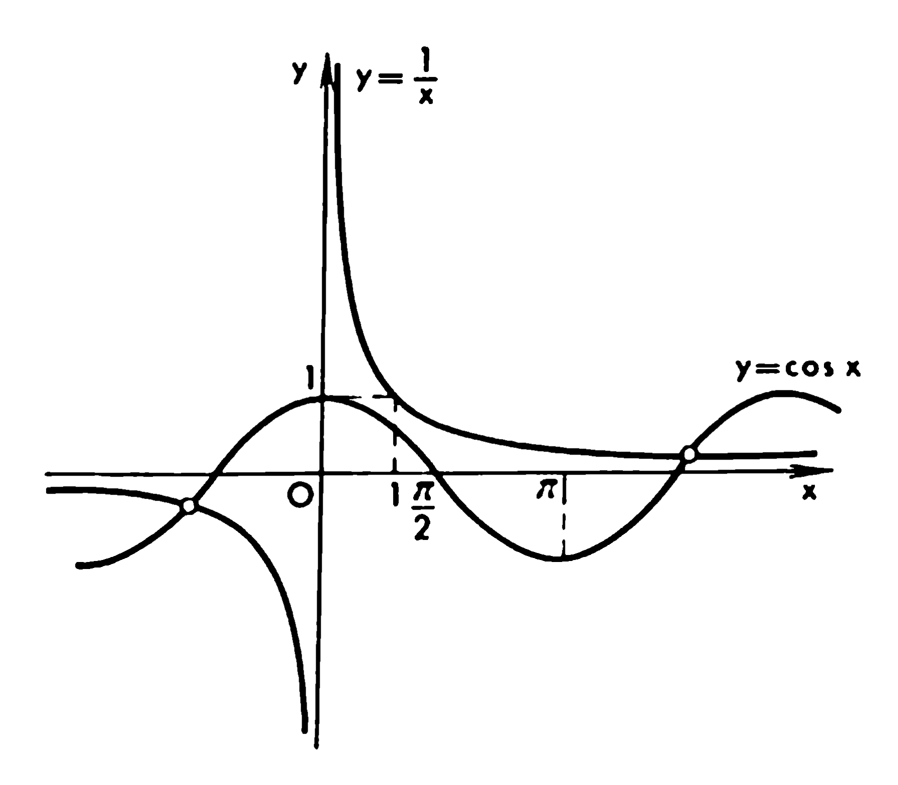


Fig. 28

forme $\pi/2 + 2k\pi$. Faisons $x - n\pi - \frac{\pi}{2} = y$ pour trouver les racines proches des points $x = \frac{\pi}{2} + n\pi$. L'équation s'écrit alors

$$y + n\pi + \frac{\pi}{2} = \frac{1}{\cos(y + n\pi + \pi/2)} = \frac{(-1)^{n+1}}{\sin y}.$$

Comme $-\frac{\pi}{2} \leqslant y \leqslant \frac{\pi}{2}$, l'équation peut s'écrire

$$y = (-1)^{n+1} \arcsin \frac{1}{y + n\pi + \frac{\pi}{2}}.$$

On a

$$\varphi(y) = (-1)^{n+1} \arcsin \frac{1}{y + n\pi + \frac{\pi}{2}}$$

et

$$\varphi'(y) = \frac{(-1)^n}{\left(y + n\pi + \frac{\pi}{2}\right)\sqrt{\left(y + n\pi + \frac{\pi}{2}\right)^2 - 1}}.$$

Il est clair qu'au voisinage du point $y = 0$ on a $|\varphi'(y)| < q < 1$ et on peut donc appliquer la méthode des approximations successives. Cherchons la solution avec une précision de 0,001 pour $n = 1$. Soit $y_0 = 0$. Il vient $y_2 \approx \varphi(y_2) \approx 0{,}204$. Donc, $y \approx 0{,}204$; $x = \frac{3}{2}\pi + y \approx 4{,}917$.

Pour trouver la première racine négative, posons $n = -1$. Nous obtenons l'équation

$$y = \arcsin \frac{1}{y - \frac{\pi}{2}}.$$

Faisons $y_0 = 0$. Alors

$$y_6 \approx \varphi(y_6) \approx -0{,}503.$$

Donc $y = -0{,}503$ et $x \approx -2{,}074$.

Pour $|n|$ grand, la méthode des approximations successives donne la formule approchée suivante:

$$y \approx \varphi(y_0) = (-1)^{n+1} \arcsin \frac{1}{n\pi + \frac{\pi}{2}} \approx \frac{(-1)^{n+1} \cdot 2}{(2n+1)\pi}.$$

Donc

$$x \approx \frac{\pi}{2}(2n+1) + \frac{(-1)^{n+1} \cdot 2}{(2n+1)\pi}.$$

12. En faisant $x_1 = 0$, il vient $x_3 \approx \varphi(x_3) \approx 1{,}088$. Donc $x \approx 1{,}088$ à 0,001 près.

13. Résolvons d'abord l'équation $x = \sqrt{\lg(x+2)}$. Nous avons $\varphi(x) = \sqrt{\lg(x+2)}$, d'où

$$\varphi'(x) = \frac{\lg e}{2(x+2)\sqrt{\lg(x+2)}}.$$

Comme $\varphi(0) = \sqrt{\lg 2} > 0$ et $\varphi(1) = \sqrt{\lg 3} < 1$, l'équation possède une solution comprise dans l'intervalle $[0, 1]$. Dans cet intervalle nous avons $|\varphi'(x)| < q < 1$. On peut donc utiliser la méthode des approximations successives. En posant $x_1 = 1$, il vient $x_5 \approx \varphi(x_5) \approx 0{,}6507$. Donc $x = 0{,}6507$ est à 0,0001 près solution de l'équation $x = \sqrt{\lg(x+2)}$.

Soit l'équation

$$x = -\sqrt{\lg(x+2)}.$$

On a

$$\varphi(x) = -\sqrt{\lg(x+2)}.$$

Comme $\varphi(0) = -\sqrt{\lg 2} = -0{,}55$ et $\varphi\left(-\frac{1}{2}\right) = -\sqrt{\lg 1{,}5} = -0{,}42$, l'équation donnée a une racine comprise dans l'intervalle $[-1/2, 0]$. En posant $x_1 = 0$, il vient $x_8 \approx \varphi(x_8) \approx -0{,}4397$. Ainsi $x = -0{,}4397$ à 0,0001 près.

14. Une racine est $x = 0$. Pour trouver la seconde racine mettons l'équation sous la forme $x = \pm\sqrt{\ln(x+1)}$. On a pour l'équation $x = \sqrt{\ln(x+1)}$

$$\varphi\left(\frac{1}{2}\right) = \sqrt{\lg\frac{3}{2}} > \frac{1}{2},$$

$$\varphi(1) = \sqrt{\ln 2} < 1.$$

L'équation a donc une racine comprise dans l'intervalle $[1/2, 1]$. Comme $\varphi(x) = \frac{1}{2(x+1)\sqrt{\ln(x+1)}}$, on a $|\varphi'(x)| < q < 1$ dans l'intervalle $[1/2, 1]$. En posant $x_1 = 1$, on obtient $x_9 \approx \varphi(x_9) \approx 0{,}7469$. Donc $x = 0{,}7469$ à 0,001 près. L'équation $x = -\sqrt{\ln(x+1)}$ ne possède pas de racines à l'exception de $x = 0$. En définitive nous avons donc deux solutions: $x = 0$ et $x = 0{,}7469$.

15. Mettons l'équation sous la forme

$$x = \sqrt{4 - \ln x}.$$

On a

$$\varphi(1) = 2, \quad \varphi(2) = \sqrt{4 - \ln 2}, \quad \varphi'(x) = \frac{-1}{2x\sqrt{4 - \ln x}}.$$

Comme $\varphi(1) = 2$, $\varphi(2) = \sqrt{4 - \ln 2} < 2$, l'équation admet une solution dans l'intervalle $[1, 2]$. Sur la fig. 29, on voit qu'il n'existe pas d'autre solution. En posant $x_1 = 2$, on obtient

$$x_4 \approx \varphi(x_4) \approx 1{,}841,$$

donc, $x = 1{,}841$ à 0,001 près.

16. Ecrivons l'équation sous la forme

$$x = 2 - \ln x.$$

On a $\varphi(x) = 2 - \ln x$, d'où $\varphi'(x) = -\frac{1}{x}$. Sur la fig. 30, on voit que l'équation admet une solution dans l'intervalle $[1, 2]$. Dans cet intervalle $|\varphi'(x)| \leqslant 1$. En faisant $x = 1{,}5$, il vient $x_{13} \approx \varphi(x_{13}) \approx 1{,}557$. Donc $x = 1{,}557$ à 0,001 près.

17. Sur la fig. 31 on voit que l'équation admet une seule racine négative. Mettons l'équation sous la forme

$$x = -\sqrt{e^x + 2}.$$

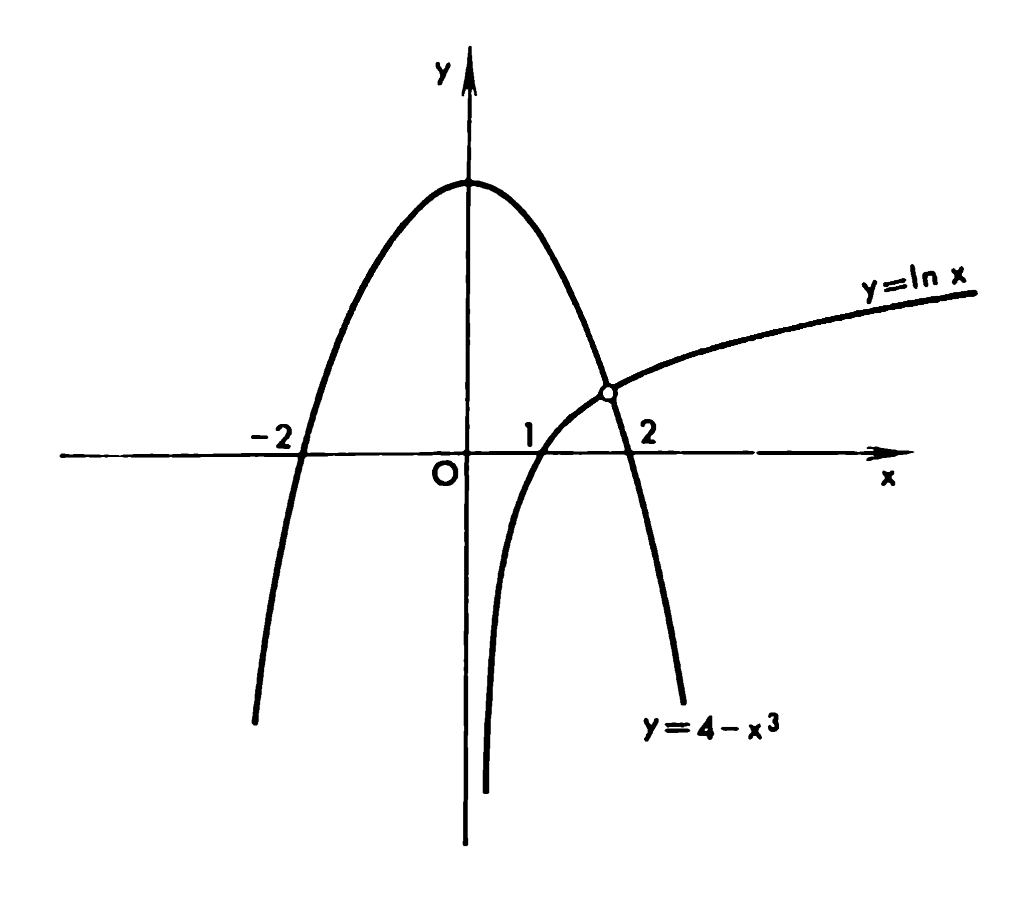


Fig. 29

Alors

$$\varphi(x) = -\sqrt{e^x+2}, \qquad \varphi'(x) = \frac{-e^x}{2\sqrt{e^x+2}}$$

et

$$\varphi(-1) = -\sqrt{2-e^{-1}} \approx -1{,}54; \qquad \varphi(-2) = -\sqrt{2+e^{-2}} \approx -1{,}46.$$

La solution est donc comprise dans l'intervalle $[-1, -2]$. En faisant $x_1 = -1$, on trouve $x_4 \approx \varphi(x_4) \approx -1{,}492$. Donc $x = -1{,}492$ à 0,001 près.

18. Il est clair qu'une solution est $x_1 = 10$. Pour trouver la seconde solution, mettons l'équation sous la forme $x = 10^{0,1x}$. On a $\varphi(x) = 10^{0,1x}$ et $\varphi'(x) = 0{,}1\cdot 10^{0,1x} \ln 10$. De plus, $\varphi(1) = 10^{0,1} >$ > 1 et $\varphi(2) = 10^{0,2} < 2$. On a donc une solution dans l'intervalle $[1, 2]$, où

$$|\varphi'(x)| \leqslant 0{,}1\cdot 10^{0,2} \ln 10 \approx 0{,}37 < 1.$$

On peut donc appliquer la méthode des approximations. En faisant $x_1 = 2$, il vient $x_7 \approx \varphi(x_7) \approx 1{,}372$. Donc $x = 1{,}372$ à 0,001 près.

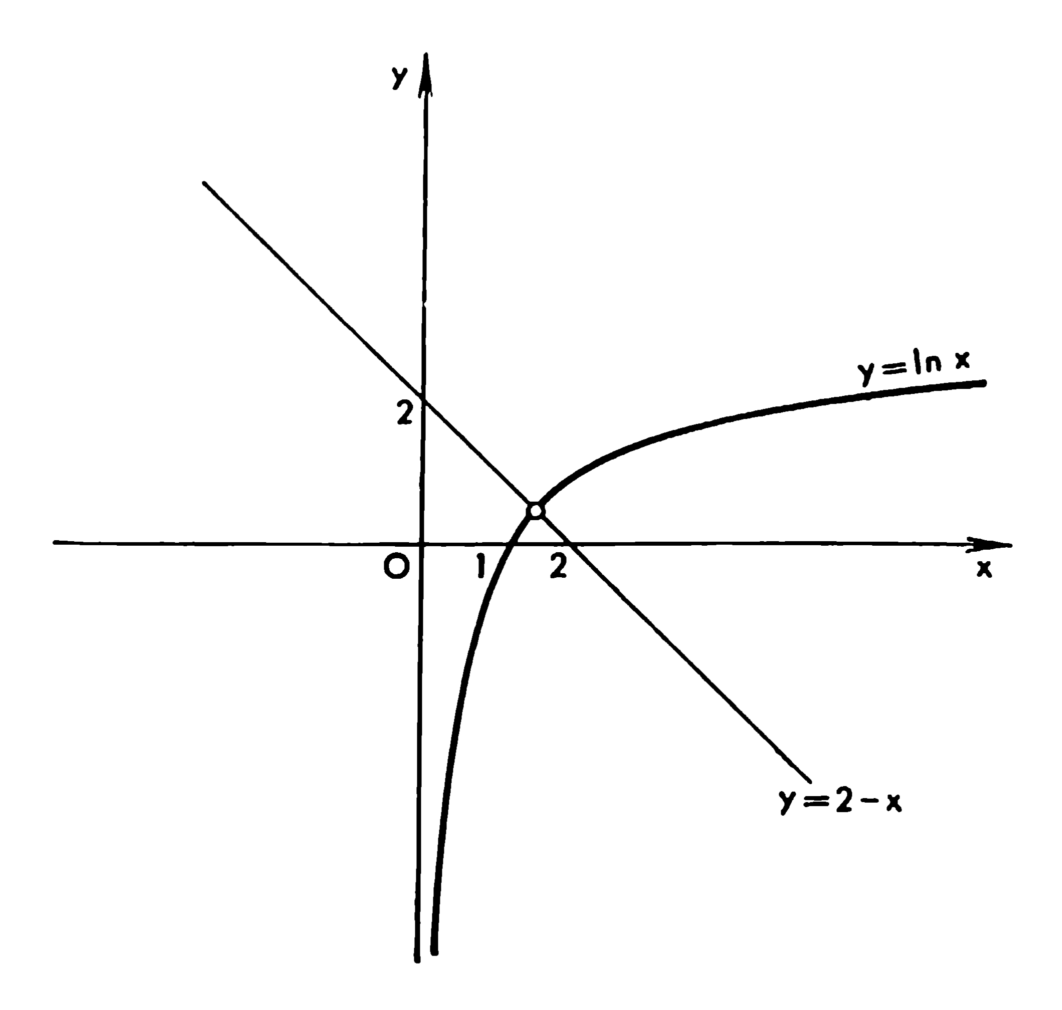


Fig. 30

19. Sur la fig. 32, on voit que l'équation a une solution dans chacun des intervalles

$$\left(\frac{\pi}{2}+n\pi,\ \frac{\pi}{2}+(n+1)\pi\right), \qquad n=0,\ 1,\ \ldots,$$

ces racines étant comprises dans la moitié de droite de chaque intervalle. Pour trouver la première racine positive, faisons le changement de variables suivant $x=\frac{3\pi}{2}-y$. L'équation s'écrit alors

$$\operatorname{cotg} y=\lg\left(\frac{3\pi}{2}-y\right),$$

d'où, puisque $0<y<\pi$,

$$y=\operatorname{arccotg}\left[\lg\left(\frac{3}{2}\pi-y\right)\right].$$

On a

$$\varphi(y)=\operatorname{arccotg}\left[\lg\left(\frac{3}{2}\pi-y\right)\right],$$

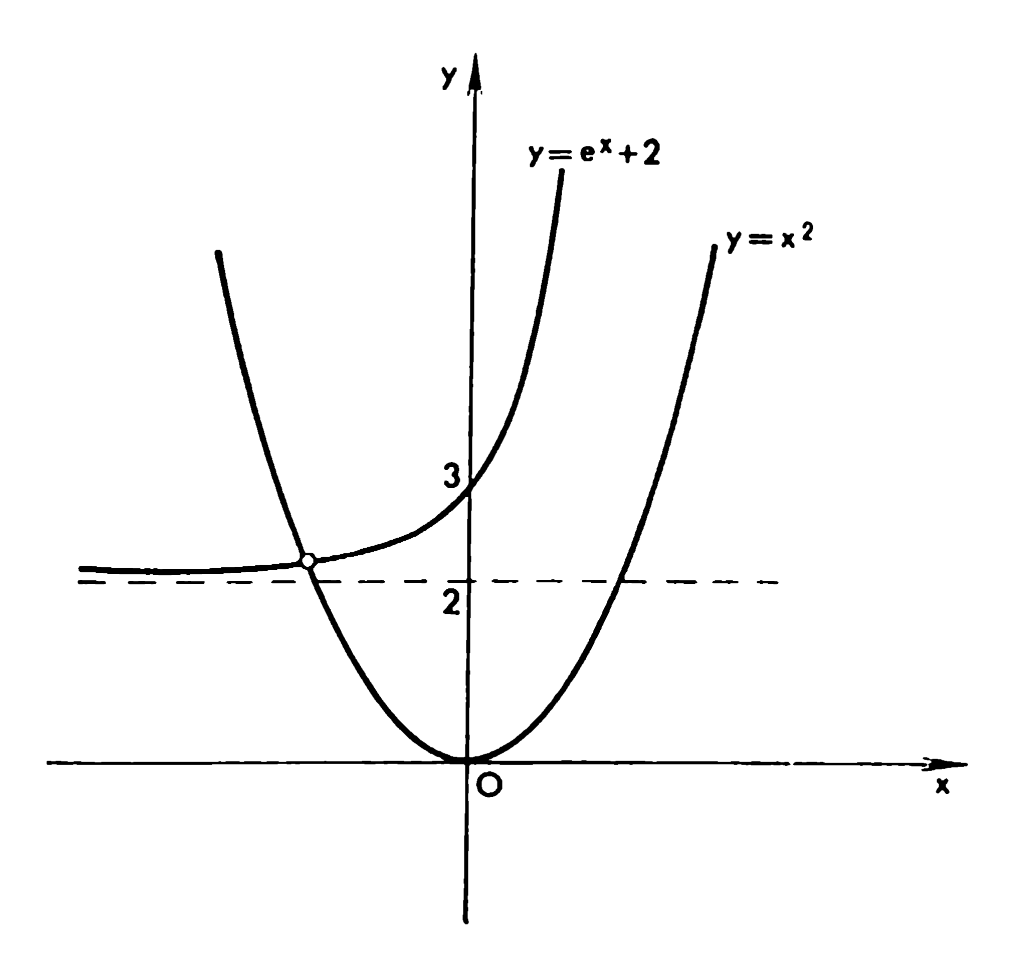


Fig. 31

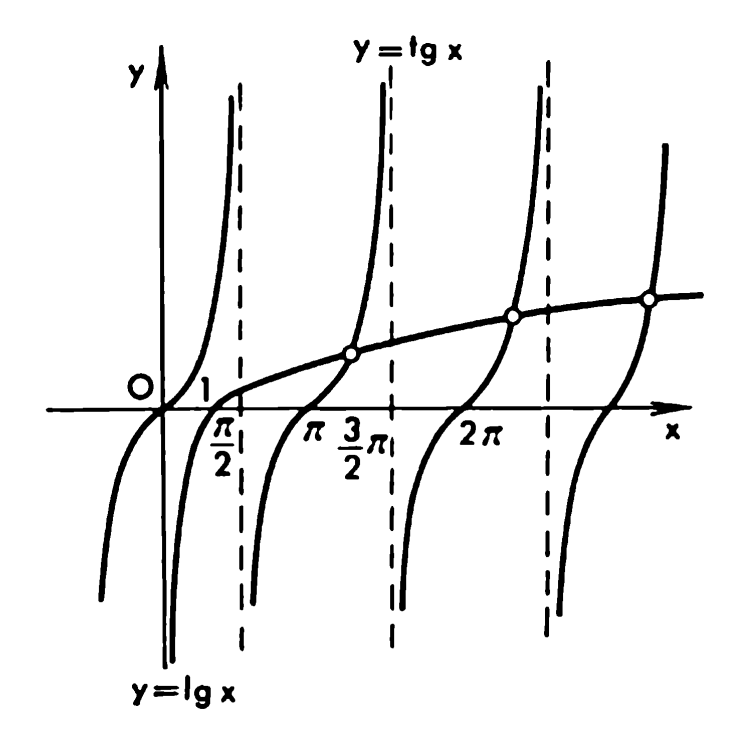


Fig. 32

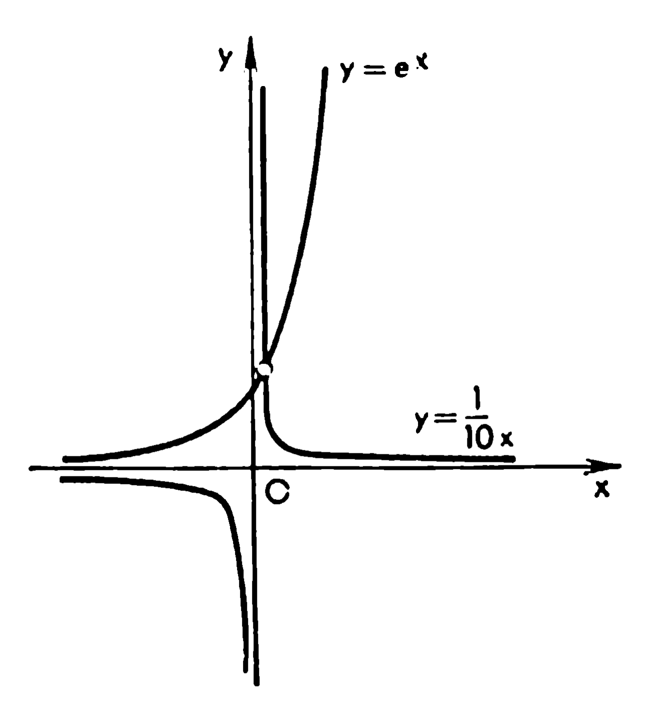


Fig. 33

et

$$\varphi'(y) = \frac{-\lg e}{\left[e + \lg\left(\frac{3}{2}\pi - y\right)\right]\left(\frac{3}{2}\pi - y\right)}.$$

Dans l'intervalle $[0, \pi]$ se trouve une racine de notre équation et comme $|\varphi'(y)| < 1$, nous pouvons nous servir de la méthode des approximations successives. En faisant $y_1 = 0$, il vient $y_4 \approx \varphi(y_4) \approx$ $\approx 1{,}059$. Donc on a $y = 1{,}059$ à 0,001 près, d'où $x = 3{,}654$.

Pour trouver la seconde racine positive, posons $x = \frac{5}{2}\pi - y$. L'équation s'écrit $y = \operatorname{arccotg}\left[\lg\left(\frac{5}{2}\pi - y\right)\right]$. En faisant $y_1 = 0$ on obtient $y_4 \approx \varphi(y_4) \approx 0{,}876$. On a donc $y = 0{,}870$ d'où $x = 6{,}984$ à 0,001 près.

20. Sur la fig. 33 on voit que l'équation admet une solution comprise entre 0 et 1. Nous avons $\varphi(x) = \frac{1}{10}e^{-x}$, $\varphi'(x) = -\frac{1}{10}e^{-x}$. Dans l'intervalle $[0, 1]$ on a $|\varphi'(x)| \leqslant \frac{1}{10}$. On peut donc appliquer la méthode des approximations successives. En faisant $x_1 = 0$, on trouve $x_4 \approx \varphi(x_4) \approx 0{,}091$. On a donc $x = 0{,}091$ à 0,001 près.

21. Soit $f(x) = x^3 - 5x + 1$. Il vient

$$f'(x) = 3x^2 - 5 \quad \text{et} \quad f''(x) = 6x.$$

D'après la formule de Newton on a

$$\beta_{n+1}=\beta_n-\frac{\beta_n^3-5\beta_n+1}{3\beta_n^2-5}.$$

Dressons le tableau de variation de la fonction:

x	-3	-2	-1	0	1	2	3
$f(x)$	-11	3	5	1	-3	-1	13

On voit que l'équation $x^3-5x+1=0$ admet des solutions dans les intervalles $[-3, -2]$, $[0, 1]$ et $[2, 3]$.

Cherchons d'abord la solution comprise dans le premier intervalle. Comme $f''(x)<0$, faisons $\beta_0=-3$ (puisque $f(\beta_0)=-11$ i.e. un nombre négatif). On a

$$\beta_1=-3-\frac{(-3)^3-5(-3)+1}{3(-3)^2-5}=-2,5.$$

On trouve ensuite $\beta_3=\beta_4\approx-2,331$. Donc la solution comprise dans l'intervalle $[-3, -2]$ est égale à $-2,331$ à 0,001 près.

Cherchons maintenant la deuxième solution. Nous avons $f''(x)>0$. On prend donc $\beta_0=0$. D'où il vient

$$\beta_1=0-\frac{0^3-5\cdot0+1}{3\cdot0^2-5}=0,2, \qquad \beta_3=\beta_4=0,202.$$

On a donc $x=0,202$ à 0,001 près. Cherchons, enfin, la troisième solution. En faisant $\beta_0=3$, on a

$$\beta_1=3-\frac{3^3-5\cdot3+1}{3\cdot3^2-5}\approx2,409.$$

On obtient ensuite $\beta_4\approx\beta_5=2,128$. Donc cette solution est égale à 2,128 à 0,001 près. En conclusion les trois solutions sont: $x_1=-2,331$, $x_2=0,202$, $x_3=2,128$.

Résolvons cette équation par la méthode améliorée des parties proportionnelles. Dans l'intervalle $[-3, -2]$ on a:

$$\alpha_1=-3-f(-3)\frac{-3-(-2)}{f(-3)-f(-2)}=-3+11\frac{(-3)}{-14}\approx-2,214.$$

Comme dans cet intervalle $f''(x) < 0$, la courbe est concave vers le bas et on tire α_2 de la relation

$$\alpha_2 = -3 - f(-3)\frac{-3-(-2{,}214)}{f(-3)-f(-2{,}214)} \approx -2{,}293.$$

On a ensuite

$$\alpha_3 = -2{,}293 - f(-2{,}293)\frac{-2{,}293+2{,}214}{f(-2{,}293)-f(-2{,}214)} \approx -2{,}331.$$

Nous retrouvons donc le même résultat avec la même précision. Les deux autres solutions se calculent de la même façon.

22. On a:

$$f(x) = x^3 - 9x^2 + 20x - 11,$$
$$f'(x) = 3x^2 - 18x + 20,$$
$$f''(x) = 6x - 18 = 6(x-3).$$

Dressons le tableau de variation de la fonction $f(x)$:

x	0	1	2	3	4	5	6
$f(x)$	−11	1	1	−5	−11	−11	1

Les solutions de cette équation sont donc comprises dans les intervalles [0, 1], [2, 3] et [5, 6].

En faisant $\beta_0 = 0$ dans l'intervalle [0, 1], il vient $\beta_4 \approx \beta_5 \approx$ $\approx 0{,}834$. Dans l'intervalle [2, 3], en faisant $\beta_0 = 3$, on trouve $\beta_2 \approx$ $\approx \beta_3 \approx 2{,}216$ et enfin on trouve $\beta_4 \approx \beta_5 \approx 5{,}249$ en prenant $\beta_1 = 0$ dans l'intervalle [5, 6]. En résumé, les trois solutions de cette équation sont (à 0,001 près):

$$x_1 = 0{,}834;\ x_2 = 2{,}216;\ x_3 = 5{,}249.$$

23. On a:

$$f(x) = x^3 - 3x^2 - 3x + 11,$$
$$f'(x) = 3x^2 - 6x - 3,$$
$$f''(x) = 6x - 6 = 6(x-1).$$

Dressons le tableau de variation de la fonction $f(x)$:

x	−2	−1	0	1	2	3
$f(x)$	−3	10	11	6	1	2

L'équation admet une racine réelle comprise dans l'intervalle [—2, —1]. Pour la trouver posons $\beta_0 = -2$. Il vient : $\beta_2 \approx \beta_3 \approx -1{,}847$. On a donc $x = -1{,}847$ à 0,001 près.

24. On a :

$$f(x) = x^5 + 5x + 1,$$
$$f'(x) = 5x^4 + 5,$$
$$f''(x) = 20x^3.$$

Dressons le tableau de variation de $f(x)$:

x	-1	0	1
$f(x)$	-5	1	7

Cette équation a donc une racine dans l'intervalle [—1, 0]. En posant $\beta_0 = -1$, nous obtenons $\beta_3 \approx \beta_4 \approx -0{,}1999$ à 0,0001 près. On a donc $x = -0{,}1999$ à 0,0001 près.

25. On a :

$$f(x) = \sin x + x - 1,$$
$$f'(x) = \cos x + 1,$$
$$f''(x) = -\sin x.$$

Dressons le tableau de variation de $f(x)$:

x	0	1	2
$f(x)$	-1	$0{,}8115$	$1{,}9093$

La racine est comprise dans l'intervalle [0, 1]. En prenant $\beta_0 = 0$, on trouve $\beta_2 = \beta_3 = 0{,}5110$. Donc $x = 0{,}5110$ à 0,0001 près.

26. On a :

$$f(x) = x^2 - 10 \lg x - 3,$$
$$f'(x) = 2x - \frac{10}{x \ln 10},$$
$$f''(x) = 2 + \frac{10}{x^2 \ln 10}.$$

Dressons le tableau de variation de $f(x)$:

x	0,5	1	2	3
$f(x)$	0,26	−2	−2,01	1,23

Les deux solutions de cette équation sont comprises l'une dans l'intervalle [0,5; 1], l'autre dans l'intervalle [2, 3]. En posant $\beta_0 = 0,5$ dans le premier, on a $\beta_2 \approx \beta_3 \approx 0,535$. Donc $x_1 = 0,535$ à 0,001 près. En faisant $\beta_0 = 3$ dans l'intervalle [2, 3], on trouve $\beta_2 \approx \beta_3 \approx 2,705$.

En résumé, les deux solutions de l'équation sont: $x_1 = 0,535$ et $x_2 = 2,705$.

27. Dans le système a) faisons $x_0 = 0$, $y_0 = 0$, $z_0 = 0$. Au bout de quelques approximations on trouve $x = 1,021$, $y = 2,150$, $z = 3,072$ à 0,001 près.

Dans le système b) faisons $x_0 = 0$, $y_0 = 0$. Au bout de quelques approximations on obtient $x = 0,520$, $y = 0,310$ à 0,001 près.

Dans le système c), en faisant $x_0 = 0$ et $y_0 = 0$, on trouve $x = 1,000$ et $y = 2,000$.

G. Chilov

Gamme simple (structure de l'échelle musicale)

Salieri ...... Pour moi,
Cela est clair, comme une simple gamme.
Pouchkine, *Mozart et Salieri*

La musique est l'art des rythmes et des sons. Au point de vue physique, le son est un processus vibratoire d'une certaine fréquence. Le la_3 correspond par exemple à des vibrations d'une fréquence de 440 hertz (vibrations à la seconde)[1]. En général, notre oreille est susceptible de percevoir des sons étalés sur une vaste bande de fréquence allant de 16 à 20 000 hertz. Jusqu'à 4000 hertz elle est capable de distinguer des sons dont les fréquences diffèrent d'un hertz seulement.

Et pourtant la musique n'emploie qu'un nombre très limité de sons. Le clavier d'un piano (fig. 1) compte en tout 90 touches blanches et noires. En les enfonçant successivement nous produisons 90 sons distincts, c'est-à-dire de hauteur différente. Voyons le tableau des fréquences de la troisième octave qui est la plus souvent utilisée dans le piano (fig. 2):

son	do_3	$ré_3$	mi_3	fa_3	sol_3	la_3	si_3	do_4
fréquence en hertz	262	294	330	349	392	440	494	523

[1]) D'après une convention internationale. Selon une légende grecque, la statue élevée au colosse Memnon aux environs de la ville de Thèbes faisait entendre ce son dès le lever du soleil, et les musiciens venaient accorder leurs instruments. Il est impossible de vérifier si cette légende est vraie, car le colosse de Memnon a cessé d'émettre ce son après des travaux de restauration effectués sous le règne de l'empereur romain Septimius Severus (193-211).

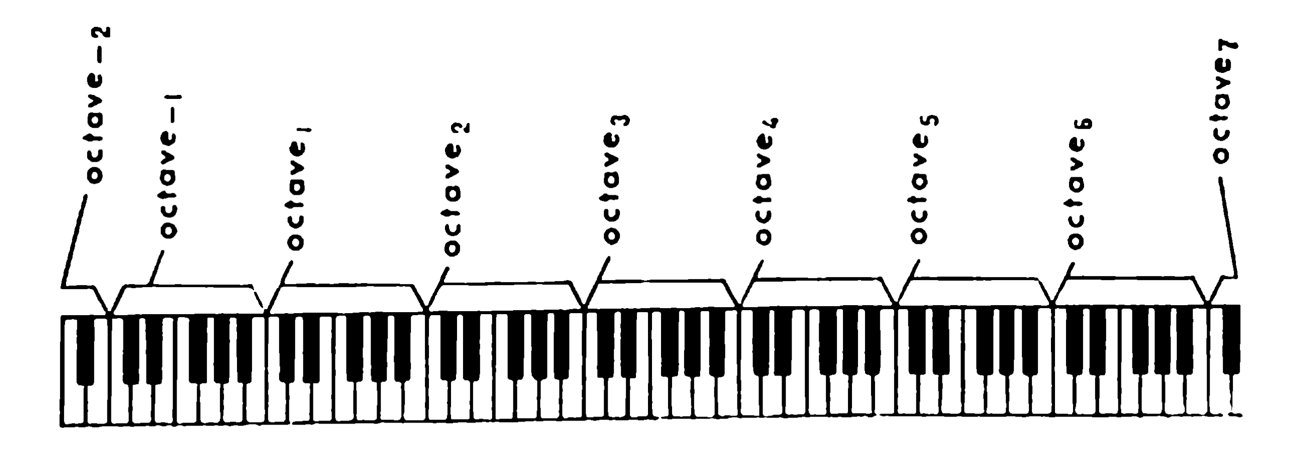


Fig. 1

Les fréquences de ces sons forment une suite pour le moins curieuse dont la seule propriété évidente est sa croissance. Par ailleurs, tous ces nombres, à l'exception de 440, sont des nombres irrationnels arrondis au plus proche entier. Ainsi, la véritable valeur de la fréquence du son mi_3 est 329,63...

Pour quelle raison a-t-on précisément posé ces sons à la base de l'échelle musicale? En fait, la question du choix de l'échelle musicale ou tonale s'est posée dans l'Antiquité déjà, mais elle n'a été définitivement résolue qu'au début du XVIIIe siècle. En vérité, cette question ne se serait jamais posée si tous les instruments de musique avaient été en mesure de produire des sons de hauteur quelconque. Si le violon, le violoncelle, etc., peuvent émettre des sons quelconques (dans les limites de leur diapason), il existe par contre des instruments tels le piano, l'orgue, la harpe, qui ne peuvent reproduire qu'un nombre limité et fixe de sons; or, l'extension de leur diapason se heurte à d'importants problèmes de construction. L'échelle musicale doit donc contenir un nombre limité de sons, mais lesquels?

Le problème n'est pas aussi simple ici que, par exemple, dans la graduation de l'échelle thermométrique où l'on repère le point de congélation et le point d'ébullition de l'eau et l'on divise ensuite l'intervalle de ces points en

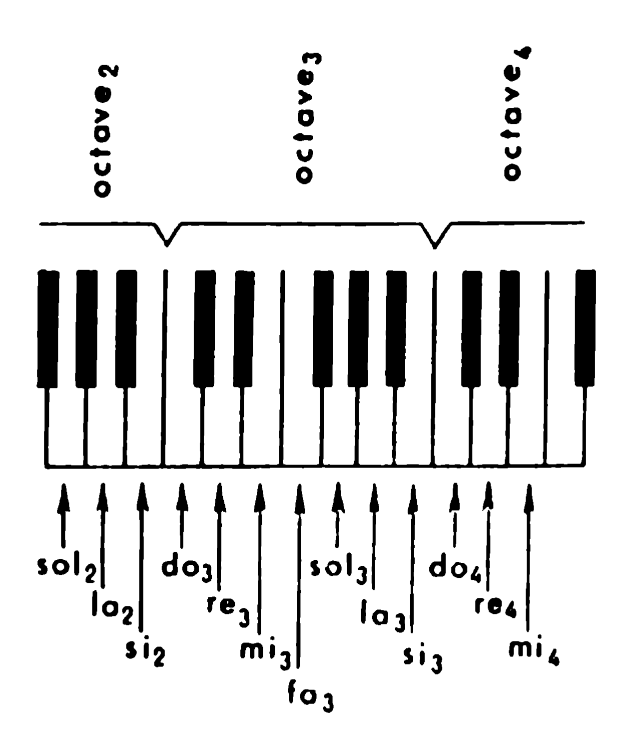


Fig. 2. *1 — do dièse = ré bémol, 2 — ré dièse = mi bémol, 3 — fa dièse = sol bémol, 4 — sol dièse = la bémol, 5 — la dièse = = si bémol.*

100 parties égales. En musique un rôle immense est celui de la consonance qui est la superposition de plusieurs sons de hauteur différente. Mais superposition n'est pas forcément consonance. Aussi, dans l'échelle musicale est-il souhaitable d'introduire avec le son fondamental des sons qui présenteront une continuité agréable pour l'oreille. Dans la suite nous expliquerons plus en détail la signification de ce que nous venons de dire.

Les instruments, tels la harpe, la lyre et la cithare, remontent à la plus haute Antiquité. Les sons y sont produits, comme d'ailleurs dans le piano, par des cordes. Rien n'interdit de penser qu'à l'Antiquité on ne soit livré sur les cordes à des expériences analogues à celle que nous allons décrire pour le piano. Les cordes du piano supportent des étouffoirs garnis de feutre qui leur interdisent de vibrer. Lorsqu'on enfonce une touche, l'étouffoir s'écarte de la corde qu'un marteau vient ensuite frapper. Si la touche reste enfoncée, la corde vibrera assez longtemps; mais dès que nous lâchons la touche, l'étouffoir vient de nouveau

en contact avec la corde et en arrête les vibrations. Si l'on enfonce lentement une touche, l'étouffoir libère la corde, mais celle-ci ne sera pas frappée par le marteau et aucun son ne sera émis. Faisons maintenant l'expérience suivante. Enfonçons lentement le la_3 qui correspond au son de fréquence 440 hertz : la corde est libérée sans émission de son. Appuyons ensuite sur le la_2 et lâchons-le aussitôt. Nous entendons le son bref du la_2 (220 hertz) qui cesse dès que la touche revient à sa position de repos. Mais ce qui est remarquable c'est le son émis par la corde du la_3 que nous avons libérée. Celle-ci commence en effet à vibrer « seule » par effet de résonance. Ceci montre que les vibrations d'une corde sont un processus bien plus compliqué qu'il ne paraît. La corde qui produit le son de fréquence 220 hertz engendre des vibrations de fréquence 440 hertz, lesquelles vibrations excitent par résonance la corde correspondante accordée sur cette dernière fréquence. On peut répéter cette expérience sur d'autres cordes de piano ou d'instruments musicaux, à chaque fois on obtiendra le même résultat : la corde de fréquence fondamentable, disons f hertz, émettra plus ou moins fortement un son de fréquence $2f$ hertz[1]). Ce phénomè-

---

[1]) On démontre, en faisant intervenir les mathématiques supérieures, que la vibration d'une corde homogène est la superposition de vibrations de fréquences f, $2f$, $3f$, . . . On peut également calculer l'amplitude de ces diverses vibrations. Si, par exemple, un marteau vient frapper une corde de longueur l' à une distance h de l'une de ses extrémités, le rapport de l'amplitude du son de fréquence $2f$ rendu à celle du son de fréquence f est égal à $\frac{1}{4}\frac{\sin(2h\pi/l)}{1+\sin(h\pi/l)}$. Le son de fréquence $2f$ n'est pas émis dans le cas seulement où la corde est excitée en son milieu (i.e. $h = l/2$). Cependant, dans ce cas (comme dans les autres d'ailleurs), les sons de fréquence $3f$, $4f$, etc., sont produits. Dans le piano, le marteau frappe les cordes au 1/8 de leur longueur. Ce point est proche de celui qui fournit le maximum de l'amplitude relative du son de fréquence $2f$, mais il n'est pas confondu avec lui car les constructeurs doivent tenir également compte des fréquences $3f$, $4f$, etc.

ne s'observe avec plus ou moins d'intensité dans les instruments à vent et à percussion (à l'exception du seul diapason qui donne un son pur).

Nous avons déjà dit que l'échelle musicale doit être étagée de telle manière que les sons ou degrés qui la composent présentent les consonances les plus harmonieuses. Deux sons dont la fréquence de l'un est le double de celle de l'autre sont très agréables à l'oreille (les vibrations de la corde se fondent les unes dans les autres et donnent un accord particulièrement agréable dit accord consonant. Seules des expériences spéciales mettent en évidence les vibrations de fréquence double).

Aussi, adopterons-nous le principe suivant : *l'échelle musicale doit outre le son de fréquence f contenir celui de fréquence $2f$.* S'il est question d'une fréquence inférieure à f, la première à figurer dans l'échelle devra être la fréquence $f/2$. On désigne par *octave* l'intervalle de deux sons tels que la fréquence de l'un soit le double de celle de l'autre. Quoique assez vaste, cet intervalle est néanmoins insuffisant. Dans la suite nous reprendrons nos expériences sur les cordes pour déterminer les sons présentant des consonances avec le son fondamental. Pour l'instant, nous allons voir une règle qui doit présider à la construction de l'échelle musicale. Plus précisément, cette échelle doit nous permettre de reproduire une mélodie plus haut ou plus bas que dans sa version originale. On sait qu'on peut chanter une chanson à des hauteurs différentes suivant le caractère de la voix de l'exécutant : le ténor dans les aiguës, la basse dans les graves. Abstraction faite de son rythme, toute mélodie est une suite d'intervalles entre les sons qui la composent. L'intervalle entre deux sons est caractérisé par le *rapport* des fréquences de ceux-ci ; nous avons vu que dans une octave ce rapport était égal à 2. Transposer un morceau de musique, c'est le jouer (ou l'écrire) soit plus haut, soit plus bas, mais en respectant le rapport des fréquences des sons dans chaque intervalle. Si nous jouons l'air du « Serin » (*mi — do — mi — do — fa — mi —*

ré) dans la troisième octave, nous utilisons les fréquences suivantes en hertz :

330—262—330—262—349—330—294.

Si nous jouons ce morceau quatre touches plus haut, nous obtenons une mélodie semblable, *si — sol — si — sol — do — si — la*. Les fréquences de ces notes sont :

494—392—494—392—523—494—440.

On vérifie aisément que les rapports des fréquences de chaque intervalle n'ont pas changé :

$$\frac{330}{262} = \frac{494}{392}, \qquad \frac{349}{262} = \frac{523}{392}, \text{ etc.}$$

Si nous modifions le rapport des fréquences, notre oreille perçoit immédiatement une altération de la mélodie. Si en particulier nous jouons la précédente mélodie une touche au-dessus, nous obtenons les sons *fa — ré — fa — ré —sol — fa — mi* de fréquences

349—294—349—294—392—349—330,

et la mélodie est dénaturée. En faisant le rapport des fréquences, on constate que

$$\frac{330}{262} \neq \frac{349}{294}, \text{ etc.}$$

En fait, on peut commencer cette mélodie sur le *fa*, mais il faudra, comme on le verra plus loin, utiliser les touches noires.

Supposons maintenant que nous ayons construit une échelle tonale satisfaisant aux deux conditions suivantes :

a) *l'échelle tonale renferme avec tout son*[1] *de fréquence* f *ceux de fréquences* $2f$ *et* $\frac{f}{2}$;

---

[1]) Pour la commodité de l'exposé nous appellerons note f_1 ou son f_1 le son de fréquence f_1.

b) *l'échelle permet de transposer une mélodie sans l'altérer.*

Soit dans l'étendue d'une octave les sons de fréquences:

$$f = f_0 < f_1 < f_2 < \ldots < f_{m-1} < f_m = 2f.$$

Ces sons forment une mélodie simple. Transposons-la d'un degré de f_0 en f_1.

La nouvelle mélodie commencera sur la note f_0 et s'achèvera sur une note f_{m+1} située à l'octave de la note f_1 (puisque $f_m = 2f_0$). Le son f_{m+1} est plus élevé que le dernier son de l'octave (f_0, f_m), mais nous affirmons qu'il vient immédiatement après f_m. En effet, si notre échelle tonale comprenait une note f' entre f_m et $f_{m+1} = 2f_1$, elle comprendrait également la note $1/2\ f'$, et de plus, l'inégalité $f_m < f' < f_{m+1}$ entraînerait:

$$f_0 < 1/2 f' < f_1.$$

Or, par hypothèse, f_1 suit *immédiatement* f_0, donc il n'existe aucune note f' entre f_m et f_{m+1}.

Après transposition d'un degré, notre mélodie doit être écrite à l'aide des notes de la même échelle tonale sur f_1 et finir sur f_{m+1}. Comme la mélodie à transposer est composée de $m + 1$ sons distincts, et que de f_1 à f_{m+1} il y a exactement $m + 1$ degrés, la nouvelle mélodie prendra la forme:

$$f_1 < f_2 < \ldots f_m < f_{m+1}.$$

Comme elle correspond à la mélodie originale, on a:

$$\frac{f_1}{f_0} = \frac{f_2}{f_1}, \quad \frac{f_2}{f_1} = \frac{f_3}{f_2}, \quad \ldots, \quad \frac{f_m}{f_{m-1}} = \frac{f_{m+1}}{f_m},$$

où

$$\frac{f_1}{f_0} = \frac{f_2}{f_1} = \ldots = \frac{f_m}{f_{m-1}} = \frac{f_{m+1}}{f_m}. \qquad (1)$$

Nous remarquons que les fréquences $f_0, f_1, \ldots, f_m$ forment une *progression géométrique*. Déterminons sa raison q. On a $f_m = q^m f_0 = 2f_0$; donc $q^m = 2$. L'échelle sera donc définie par son nombre m de degrés.

Pour plus de commodité passons des fréquences $f_0, f_1, \ldots$ aux logarithmes binaires $\log_2 f_0$, $\log_2 f_1, \ldots$ L'octave $(f_0, 2f_0)$ se transforme en l'intervalle de longueur 1 :$(\log_2 f_0$, $\log_2 2f_0 = \log_2 f_0 + 1)$, et la progression géométrique $f_0, f_1, \ldots, \ldots, f_m$ en la progression arithmétique $\log_2 f_0, \log_2 f_1, \ldots \ldots, \log_2 f_m$ de raison $\log_2 \sqrt[m]{2} = \frac{1}{m}$. Donc notre échelle tonale sera représentée sur l'axe des logarithmes par les points A, $A + \frac{1}{m}$, $A + \frac{2}{m}$, $\ldots$, $A + 1$ où $A = \log_2 f_0$.

A quel critère doit obéir le choix de m?

Reprenons nos expériences sur la corde pour déterminer les sons qu'elle émet en vibrant. Voyons si outre les sons de fréquence f et $2f$ elle émet celui de fréquence $3f$. Pour cela utilisons le la_2 de fréquence 220 hertz, et le mi_4 de fréquence 660 hertz. Enfonçons lentement le mi_4 de manière à libérer la corde; puis, comme précédemment, appuyons sèchement sur le la_2. Aussitôt *nous entendons vibrer la corde du* mi_2. Donc, le *son de fréquence triple est également produit*. En poursuivant cette expérience, on constaterait la présence des sons de fréquence quadruple (ce qui n'a rien d'étonnant puisque $4f = 2 \cdot 2f$), quintuple, etc. Cependant, les sons de fréquence triple et plus sont nettement moins exprimés. Les sons de fréquence $2f$, $3f$, $4f$, $\ldots$ sont appelés *harmoniques* du son principal ou fondamental de fréquence f; la superposition d'un son et de ses harmoniques caractérise le timbre. Celui-ci permet de distinguer deux sons identiques émis à l'unisson par des instruments différents. Le timbre de la voix dépend du nombre et des intensités relatives des harmoniques. Un son simple sans harmonique est terne, aussi n'est-il pas utilisé en musique.

Introduisons maintenant la condition suivante:

c) *l'échelle tonale doit contenir le son de fréquence f et celui de fréquence $3f$.*

Puisque nous avons convenu précédemment que l'échelle tonale devait contenir le son de fréquence $1/2\ f$, nous

remarquons qu'elle doit aussi contenir le son de fréquence $\frac{1}{2} 3f = \frac{3}{2} f$. Cette dernière fréquence nous intéresse particulièrement parce qu'elle est comprise dans l'intervalle $(f, 2f)$, qui représente l'étendue de notre octave. Ainsi l'octave $(f_0, 2f_0)$ doit également contenir le son de fréquence $3/2 f_0$. Le logarithme de la fréquence du k-ième degré est $A + \frac{k}{m}$, celui de la fréquence $3/2\, f_0$ est $A = \log_2 3/2$.

D'où l'équation:

$$\log_2 \frac{3}{2} = \frac{k}{m}, \qquad (2)$$

qui doit être vérifiée quels que soient k et m entiers. On s'assure facilement que *cette équation n'est pas résoluble dans les entiers*; autrement dit, $\log_2 3/2$ est un nombre *irrationnel*. En effet, la définition (2) du logarithme entraîne

$$2^{\frac{k}{m}} = \frac{3}{2},$$

où, en élevant à la puissance m,

$$2^k = \left(\frac{3}{2}\right)^m \quad \text{et} \quad 2^{k+m} = 3^m.$$

Or, le premier membre de cette égalité est pair quels que soient k et m, tandis que le second membre est impair.

Cette contradiction montre qu'une *échelle tonale logarithmique uniforme ne peut à la fois contenir le son de fréquence $3/2\, f$ et celui de fréquence f*. L'intervalle $(f, 3/2f)$ est appelé *quinte juste*. Nous voyons immédiatement que les quintes justes n'existent pas dans l'échelle tonale logarithmique uniforme.

Donc il faut soit renoncer à l'uniformité de l'échelle tonale, soit aux quintes justes. L'uniformité de l'échelle est indispensable pour la transposition d'une mélodie sans altération, et il serait fâcheux d'y renoncer. Force est donc de se séparer des quintes justes: essayons de former une suite

de nombres rationnels $\frac{k}{m}$ approchant le nombre irrationnel $\log_2 \frac{3}{2}$ et tels que les fréquences qui leur correspondent présentent une différence inférieure à 1 hertz, qui ne pourra être perçue par notre oreille. Voyons quelle doit être la précision des calculs. La première octave qui va de 262 à 523 hertz s'étend donc sur 260 hertz ; sur l'échelle logarithmique cet intervalle est de longueur 1. Donc 1 hertz correspond approximativement à une division de 0,004 sur l'échelle logarithmique. Les nombres $\frac{k}{m}$ et le nombre $\log_2 \frac{3}{2} =$ $= 0{,}585\ldots$ doivent présenter une différence inférieure à la moitié de la deuxième décimale.

Outre l'intervalle de quinte $3/2f$, il en existe d'autres dont les plus fréquents, ainsi qu'il ressort d'une analyse de la musique populaire[1]) sont : l'intervalle d'octave 2, de quinte 3/2, de tierce 5/4, de quarte 4/3, de sixte 5/3, de seconde 9/8 et de septième 15/8. Ecrivons les valeurs correspondantes des logarithmes binaires :

$$\log_2 2 = 1,\ \log_2 3/2 = 0{,}585,\ \log_2 4/3 - 0{,}416,$$
$$\log_2\ 5/3 = 0{,}737,\quad \log_2 5/4 = 0{,}323,\ \log_2\ 9/8 = 0{,}169,$$
$$\log_2\ 15/8$$[2]).

---

[1]) On obtient les mêmes rapports à partir de considérations théoriques. La tierce 5/4 est le représentant du son fondamental quintuple, la quarte 4/3 est à une quinte plus bas que le 2, la sixte 5/3, à une quinte plus bas que $5/2 = 2 \cdot 5/4$, la seconde $9/8 = 1/2 \cdot 3/2 \cdot 3/2$ est le représentant de la double quinte; enfin, la septième 15/8 est à une quinte plus haut que la tierce.

[2]) Si on n'a pas de tables de logarithmes binaires sous la main, on peut se servir des logarithmes décimaux. Soit $2^{\log_2 x} = x$. On a

$$\log_2 x \cdot \lg 2 = \lg x,$$

d'où

$$\log_2 x = \frac{\lg x}{\lg 2}.$$

Il nous faut construire une échelle uniforme dont les degrés soient de préférence proches de ces nombres et tels que le plus grand d'entre eux soit $\log_2 3/2 = 0{,}585$ comme correspondant à l'intervalle le plus naturel dans les limites de l'octave.

La *fraction continue*, i.e. une fraction de la forme

$$\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3\ldots}}}, \tag{3}$$

où $a_1, a_2, \ldots$ sont des entiers naturels, est un excellent moyen pour approcher tout nombre irrationnel par des nombres rationnels.

On sait que tout nombre α compris dans l'intervalle $[0, 1]$ peut être représenté par une fraction continue (infinie si α est irrationnel). Les expressions rationnelles

$$\frac{1}{a_1}, \quad \cfrac{1}{a_1+\cfrac{1}{a_2}}, \quad \cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3}}}, \quad \text{etc.,}$$

sont appelées *réduites* de la fraction continue (3). Entre la fraction

$$\cfrac{1}{a_1+\cfrac{1}{a_2+\ldots+\cfrac{1}{a_3}}}=\frac{p_n}{q_n}$$

et le nombre α nous avons les relations:

$$\left|\frac{p_n}{q_n}-\alpha\right| \leqslant \frac{1}{q_n^2},$$

$$\left|\frac{p_n}{q_n}-\alpha\right| \leqslant \frac{p}{q},$$

où p est quelconque et $q \leqslant q_n$.

Cherchons les premières réduites de la fraction continue correspondant au nombre $x = \log_2 3/2$. D'après la définition du logarithme, on a

$$2^x = \frac{3}{2}. \qquad (4)$$

Comme $x < 1$, en substituant $y = 1/x$, il vient $y > 1$. L'équation (4) s'écrit alors

$$\left(\frac{3}{2}\right)^y = 2. \qquad (5)$$

Il est évident que y est compris entre 1 et 2 (puisque $(3/2)^1 = 3/2 < 2$, $(3/2)^2 = 9/4 > 2$. Posons maintenant $y = 1 + 1/z$; on remarque que $1/z < 1$ et $z > 1$. L'équation (5) s'écrit

$$\frac{3}{2}\left(\frac{3}{2}\right)^{\frac{1}{z}} = 2, \qquad \left(\frac{3}{2}\right)^{\frac{1}{z}} = \frac{4}{3},$$

d'où

$$\left(\frac{4}{3}\right)^z = \frac{3}{2}. \qquad (6)$$

On voit que l'inconnue z est comprise entre 1 et 2 ($4/3 < 3/2$, $(4/3)^2 = 16/9 > 3/2$). Posons $z = 1 + 1/u$; l'équation (6) prend la forme

$$\frac{4}{3}\left(\frac{4}{3}\right)^{\frac{1}{u}} = \frac{3}{2}, \qquad \left(\frac{4}{3}\right)^{\frac{1}{u}} = \frac{9}{8},$$

d'où

$$\left(\frac{9}{8}\right)^u = \frac{4}{3}. \qquad (7)$$

On voit que u est compris entre 2 et 3 ($(9/8)^2 = 81/64 < 4/3$, $(9/8)^2 = 729/512 > 4/3$), aussi poserons-nous $u = 2 + 1/v$ où $v > 1$. Nous obtenons:

$$\left(\frac{9}{8}\right)^2\left(\frac{9}{8}\right)^{\frac{1}{v}} = \frac{4}{3}, \quad \text{d'où} \quad \left(\frac{9}{8}\right)^{\frac{1}{v}} = \frac{256}{243},$$

ou encore

$$\left(\frac{256}{243}\right)^v = \frac{9}{8}. \qquad (8)$$

Il est maintenant plus commode de poursuivre les calculs en se servant de la table des logarithmes. Prenons le logarithme de (8):

$$v\,[\lg 256 - \lg 243] = \lg 9 - \lg 8,$$

ou

$$v\,[2{,}4082 - 2{,}3856] - 0{,}9542 - 0{,}9031$$

et

$$0{,}0226v = 0{,}0511.$$

Il est clair que v est compris entre 2 et 3. Les calculs seraient indéfinis, mais nous préférons nous arrêter là. Finalement nous obtenons:

$$x = \frac{1}{y} = \frac{1}{1+\frac{1}{z}} = \frac{1}{1+\frac{1}{1+\frac{1}{u}}} = \frac{1}{1+\frac{1}{1+\frac{1}{2+\frac{1}{v}}}} =$$

$$= \frac{1}{1+\frac{1}{1+\frac{1}{2+\frac{1}{2+\dots}}}},$$

ce qui nous donne les premiers termes de la fraction continue cherchée.

Les réduites correspondantes sont donc:

$$\frac{1}{1} = 1; \qquad \frac{1}{1+\frac{1}{1}} = \frac{1}{2},$$

$$\frac{1}{1+\frac{1}{1+\frac{1}{2}}} = \frac{3}{5}; \qquad \frac{1}{1+\frac{1}{1+\frac{1}{2+\frac{1}{2}}}} = \frac{7}{12}.$$

Les deux premières approchent grossièrement le nombre cherché. La troisième $\frac{k}{m} = \frac{3}{5} = 0,600$ présente déjà une erreur relativement petite (0,015) par rapport à la grandeur $\log_2 3/2 = 0,585$ qui nous intéresse. Cependant, cette erreur est 4 fois supérieure à l'erreur 0,004 désirée. Si d'autre part nous considérons l'échelle composée des multiples de 1/5, i.e. 1/5, 2/5, 3/5, 4/5, 1, nous constatons que certains nombres nous intéressant, notamment $\log_2 5/3 = 0,727$ et $\log_2 9/8 = 0,169$, sont éloignés de ces multiples.

Passons à la dernière approximation, $\frac{k}{m} = \frac{7}{12} = 0,583$. Elle est suffisamment proche de la valeur cherchée 0,585 puisque l'erreur 0,002 n'est que la moitié de l'erreur admise. L'échelle tonale se construit sur l'axe logarithmique dont l'intervalle [0, 1] est divisé en douze parties égales par les points :

$$\frac{1}{12} = 0,083, \quad \frac{2}{12} = 0,167, \quad \frac{3}{12} = 0,250, \quad \frac{4}{12} = 0,333,$$

$$\frac{5}{12} = 0,418, \quad \frac{6}{12} = 0,500, \quad \frac{7}{12} = 0,583, \quad \frac{8}{12} = 0,667,$$

$$\frac{9}{12} = 0,750, \quad \frac{10}{12} = 0,833, \quad \frac{11}{12} = 0,917, \quad \frac{12}{12} = 1,000,$$

dont le septième est très proche de la quinte.

Nous remarquons que les valeurs $\log_2 \frac{4}{3} = 0,416$, $\log_2 \frac{5}{3} = 0,737$, $\log_2 \frac{5}{4} = 0,323$, $\log_2 \frac{9}{8} = 0,169$, $\log_2 \frac{15}{8} = 0,908$ qui nous intéressent sont proches des points graduateurs de l'échelle (fig. 3), bien que la précision ne soit pas la même que pour $\log_2 \frac{3}{2}$.

Donc la division de l'échelle tonale en douze parties égales est celle qui nous convient le mieux.

Maintenant nous sommes en mesure d'expliquer totalement la loi régissant les fréquences de l'octave. Tout d'abord

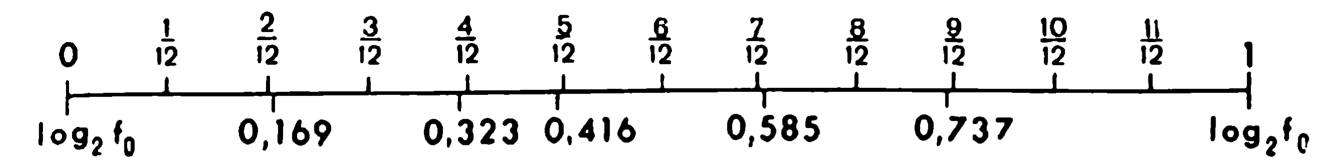


Fig 3

nous allons fixer chaque degré de cette échelle en convenant que le rapport de deux fréquences successives est constant et égal à $\sqrt[12]{2}$.

On appelle *demi-ton* l'intervalle séparant deux degrés consécutifs de cette échelle. On appelle *ton* l'intervalle composé de deux demi-tons. L'octave se divise en 6 tons ou 12 demi-tons. Les sons fondamentaux de l'octave sont voisins des sons de fréquence f_1, f_2, . . . représentés sur la fig. 3, de sorte qu'au lieu du son de fréquence f_1 on peut prendre le plus proche degré, i.e. $\frac{2}{12}$, au lieu de f_2 le degré $\frac{4}{12}$, etc. (fig. 4). Si la première note de l'octave est *do*, la deuxième note, séparée de la première par un ton, s'appelle *ré*; la troisième, séparée de la deuxième par un ton, s'appelle *mi*; et la quatrième qui est séparée de la troisième par un demi-ton s'appelle *fa*. Ces quatre notes fondamentales constituent un tétracorde (terme très ancien inspiré aux Grecs par leur lyre à quatre cordes). L'autre moitié de l'octave est constituée par un second tétracorde ou tétracorde supérieur, identique au premier et commençant sur la note *sol* qui correspond au son de fréquence $\sim 3/2 f_0$. A un ton vient le *la*, puis le *si* et enfin à un demi-ton le do_4, qui est à l'octave du do_3. La juxtaposition de ces deux tétracordes donne la première octave que nous avons étudiée plus haut. En effet, les logarithmes binaires des fréquences indiquées dans

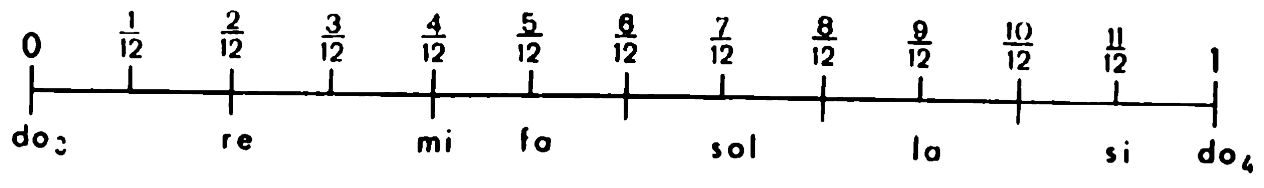


Fig. 4

la table sont :

Note	*do*$_3$	*ré*$_3$	*mi*$_3$	*fa*$_3$	*sol*$_3$	*la*$_3$	*si*$_3$	*do*$_4$
fréquence f en hertz	262	294	330	349	392	440	494	523
$\log_2 f$	8,031	8,198	8,365	8,448	8,615	8,781	8,948	9,031
$\log_2 \frac{f}{f_0} = \log_2 f - \log_2 f_0$	0	0,167	0,334	0,417	0,584	0,750	0,917	1,000

Nous voyons que la différence des logarithmes représente précisément les nombres qui figurent dans notre échelle tonale. La structure de la troisième octave est donc définie.

Outre les sept notes fondamentales, l'octave comporte cinq notes auxiliaires qui superposées aux premières donnent une échelle à douze degrés. Elles sont désignées par les notes fondamentales affectées d'un *dièse* qui les élevera d'un demi-ton ou d'un *bémol* qui les abaissera d'un demi-ton. Ainsi le *do dièse* et le *ré bémol* représentent le même son. Les cinq notes auxiliaires correspondent aux touches noires de la première octave (voir fig. 2).

Si une mélodie est composée exclusivement des sons fondamentaux de la gamme, elle n'est jouée qu'avec les touches blanches (par exemple le « Serin » commence sur le *mi*). Mais si nous voulons transposer le « Serin » d'un demiton plus haut, il nous faudra utiliser les touches noires puisque la deuxième note de la mélodie qui est un *do* devient *ré bémol*.

La gamme : *do* ou ut — *ré* — *mi* — *fa* — *sol* — *la* — *si* — *do*, qui est constituée de deux tons, un demi-ton, trois tons, et un demi-ton, est appelée gamme naturelle

majeure ou gamme d'*ut majeur.* Chaque note de cette échelle est le point de départ d'une nouvelle gamme qui portera le nom de la note sur laquelle elle commencera. Ainsi, pour former la gamme *ré majeur* on diésera le *fa* et le *do* pour obtenir la structure de la gamme majeure c'est-à-dire : *ré — mi — fa dièse — sol — la — si — do dièse — ré.* La gamme *ré majeure* sera donc *ré — mi — fa dièse — sol — la — si — do dièse — ré.* On formerait ainsi 11 gammes majeures.

Il existe encore des gammes de sept degrés, mais qui présentent un autre rapport entre les intervalles. Ainsi la gamme naturelle d'*ut mineur* : *do — ré — mi bémol — fa — sol — la bémol — si bémol — do*, est constituée d'un ton, un demi-ton, deux tons, un demi-ton, deux tons. Les trois degrés principaux de la gamme d'*ut majeur* sont la tonique *do*, la médiante *mi* et la dominante *sol.* Leur réunion forme l'accord parfait de tonique ou accord parfait majeur à quelques octaves près (la fréquence de sol_2 est le triple de celle de do_1 et celle de mi_3 le quintuple de do_1). C'est ce qui explique sans doute que cet accord soit nettement et pleinement perçu. Les trois sons principaux de la gamme d'*ut mineur* sont *ut — mi bémol — sol.* Ils constituent un accord parfait mineur que l'on obtient en bémolisant la deuxième note d'un accord parfait majeur, d'où une dissonance latente entre le *mi bémol* et le mi_5 qui est un harmomique du do_3 ; ceci explique sans doute la saveur particulière des accords mineurs. Les œuvres musicales portent souvent l'indication du mode, majeur ou mineur, dans lequel elles sont composées : exemple, ballade de Chopin en sol mineur, polonaise en *la bémol majeur.*

La création de l'échelle tonale uniforme (i.e. divisée en douze intervalles égaux) est le résultat de la longue évolution de la musique et des mathématiques. On comprend qu'elle n'ait pu apparaître avant l'élaboration de la théorie des nombres irrationnels et des logarithmes. Or, les mathématiciens n'ont commencé à user librement de ces notions

qu'au XVII siècle. Vers 1700, le savant et musicien allemand Andreas Werckmeister proposa l'échelle que nous venons de décrire et construisit même un piano qu'il accorda en conséquence. Jusque-là les instruments musicaux étaient accordés suivant le principe des intervalles justes (quinte, tierce, etc.), ce qui inévitablement posait des problèmes délicats dans l'utilisation d'autres tonalités et dans la modulation (c'est-à-dire le passage d'une tonalité à une autre), et limitait par là même l'essor de la musique. Les musiciens n'adoptèrent pas immédiatement l'échelle de Werckmeister; Diderot dont l'esprit vivace s'intéressait également à la musique fut au rang de ses adversaires, estimant qu'une échelle sans intervalles justes ne pouvait servir de fondement à la musique. Mais le célèbre compositeur allemand Jean Sébastien Bach démontra la vitalité de ce nouveau système qu'il adopta pour écrire deux tomes groupés sous le titre « Le clavecin bien tempéré » (1722-1744). Chacun de ces tomes comprenait 24 pièces (des préludes et des fugues) composées dans les 12 tonalités majeures et les 12 tonalités mineures à raison d'une pièce par tonalité. Les œuvres de Bach firent époque dans l'évolution de la musique. Ce nouveau système ne tarda pas à être adopté par d'autres compositeurs. Les perspectives qu'il ouvrit sont telles que même de nos jours ses possiblités sont loin d'être épuisées. Les altérations des intervalles « populaires » justes ne sont perçues que par une oreille bien exercée, cependant, ce « défaut » est largement compensé par le libre choix des tonalités et la facilité des modulations. Au XXe siècle des compositeurs ont proposé de porter à 24, 48 ou 53 le nombre des degrés de l'octave afin d'obtenir des intervalles proches des intervalles justes. On a même construit des instruments spéciaux, mais ceux-ci n'ont pas trouvé audience.

En conclusion signalons une circonstance à laquelle la science musicale n'a pas encore trouvé d'explication théorique. En vertu de notre construction, les 12 tonalités

majeures, tout comme les 12 tonalités mineures, doivent être identiques du point de vue de leur sonorité. Pourtant, les musiciens estiment que les tonalités sont douées de traits particuliers. On admet généralement que le *do majeur* crée une ambiance sereine, radieuse (la sonate « Aurore » de Beethoven), le *mi majeur* est réservé pour les sentiments passionnés, agités (« Cantique d'amour » de Liszt) ; le *fa dièse majeur* pour les sentiments sublimes (« Au printemps » de Grieg) ; le *do majeur* pour une affliction extrême (La marche funèbre » de la symphonie héroïque de Beethoven) ; le *mi bémol mineur* pour les états d'âme tragiques (la romance de Pauline de « La Dame de pique » de Tchaïkovski). On ne sait pas encore si ces conventions résultent de lois objectives, ou si elles sont le fait de la tradition. Il se peut au reste que l'accordement des instruments de musique donne, en raison des particularités de l'oreille, non pas des intervalles continus, un peu rigides même, mais des intervalles légèrement atténués, de sorte que par exemple le rapport des fréquences de l'intervalle *do — sol* n'est pas exactement le même que celui des fréquences de l'intervalle *mi — si*, comme le laissait prévoir un accordement idéal. Quoi qu'il en soit, la science est en perpétuel progrès et tôt ou tard elle fournira une explication à tous les points sombres de la musique.

V. Ouspenski

Quelques applications de la mécanique aux mathématiques

§ 1 PROBLÈME DE LA TANGENTE À UN CERCLE

On sait que *la tangente* à un cercle est par définition une droite ne possédant qu'un point commun, appelé *point de contact*, avec ce cercle. En géométrie, on démontre que la tangente est perpendiculaire au rayon passant par le point de contact. Donnons une démonstration de ce théorème, basée exclusivement sur des notions de mécanique.

A ces fins faisons l'expérience suivante. Imaginons-nous une masse m suspendue à l'extrémité M d'un fil dont l'autre extrémité O est fixe. On sait empiriquement que

(A) *La masse m occupe sa position la plus basse.*

Cette constatation qui est capitale pour les raisonnements ultérieurs est l'évidence même; toutefois, il est utile de souligner qu'elle est une manifestation particulière d'une loi générale (formulée dans le « principe du minimum de l'énergie potentielle») que nous retrouverons dans une situation plus complexe au § 4. Plus exactement, le « principe du minimum de l'énergie potentielle » permet de déduire de la proposition (A) la proposition évidente (E) qui suit:

(E) *La masse m ne possède qu'un seul état d'équilibre*, i.e. au repos la position de la masse m est définie par la position du point O et par la longueur du fil.

L'unicité de l'état d'équilibre (proposition E) entraîne que

(C) *La masse m se trouve sur la verticale passant par le point fixe O.*

En effet, si la masse ne se trouvait pas sur la verticale, on obtiendrait une nouvelle position d'équilibre (fig. 1) en faisant pivoter cette masse autour de la verticale. Ce raisonnement repose donc sur l'observation suivante:

(S) *Si un corps, ou un système de corps, se trouvant en état d'équilibre pivote autour de la verticale, son nouvel état sera un état d'équilibre.*

Les propositions (A) et (C) entraînent que

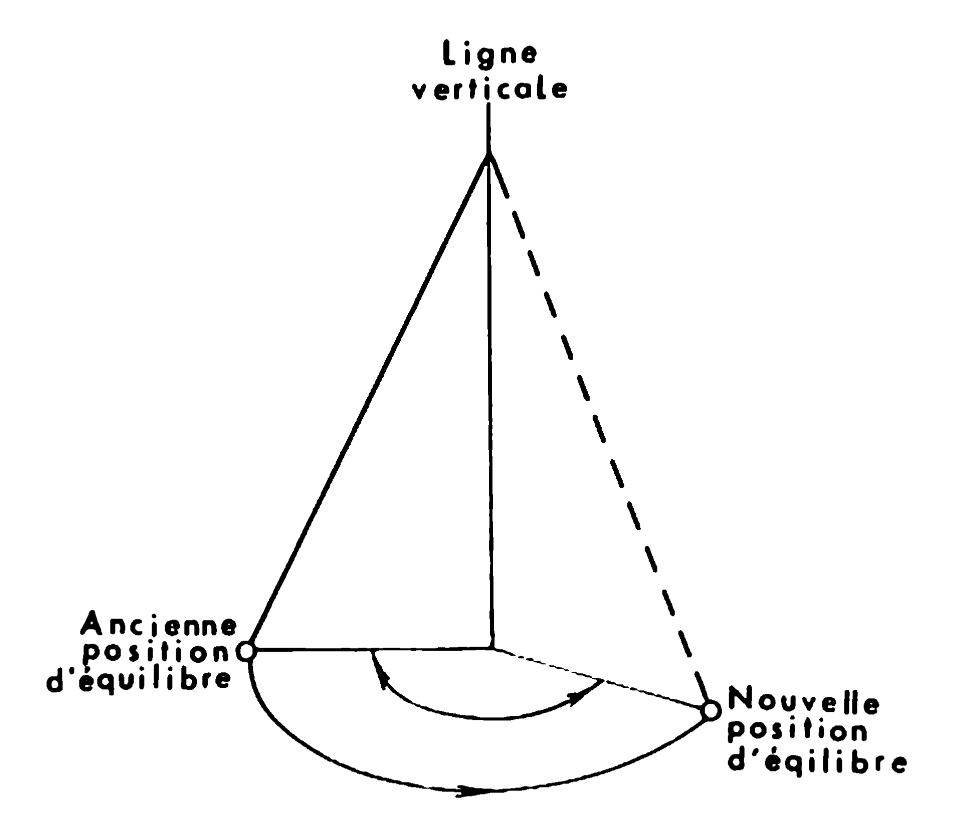


Fig. 1

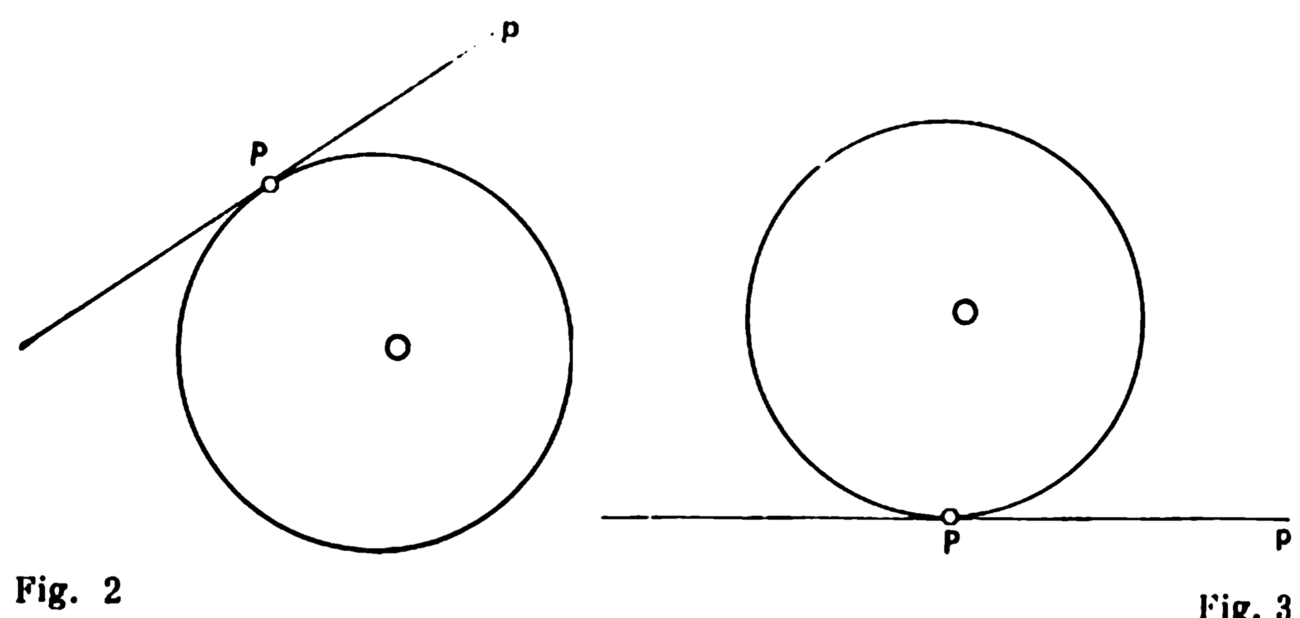


Fig. 2

Fig. 3

(B) *La distance de la masse m au point fixe O est égale à la longueur du fil MO* (ce qui signifie que le fil est tendu).

(Remarquons au passage que les propositions (B) et (C) sont aussi évidentes que celles qui les ont entraînées).

Passons à présent à la démonstration de notre théorème. Soit un cercle de centre O et une droite p tangente à ce cercle en P (fig. 2). Il nous faut prouver que $OP \perp p$. Pour cela portons-nous sur la fig. 3, i.e. nous supposerons que la droite p est horizontale et le cercle situé au-dessus d'elle. Le point P sera donc le plus bas point du cercle. Prenons à présent un fil de longueur égale au rayon du cercle et fixons l'une de ses extrémités au centre O. A son autre extrémité M suspendons une masse m. Ceci posé, nous allons prouver que M est confondu avec le point P. En effet, d'abord en vertu de (B) l'extrémité M ne peut se trouver que sur le cercle, ensuite en vertu de (A) cette extrémité se trouve au point le plus bas du cercle, i.e. en P. Donc, le fil sera tendu suivant le rayon OP. Il résulte alors de (C) que ce rayon est perpendiculaire à la droite p, ce qu'il fallait démontrer.

L'exemple étudié n'est sans doute pas très instructif dans la mesure où il démontre un théorème bien connu et relativement simple. Cependant, comme nous le verrons dans la suite, cette méthode sera appliquée à la démonstration de nouveaux théorèmes. Nous aurons alors à nous servir des éléments exposés dans ce paragraphe pour démontrer « mécaniquement » le théorème de la tangente à un cercle. Nous allons tout d'abord généraliser ce théorème, i.e. examiner le théorème de la tangente à une ellipse.

§ 2 PROBLÈME DE LA TANGENTE À UNE ELLIPSE

Le cercle est par définition le lieu géométrique des points A situés à une égale distance l d'un point fixe O appelé centre. Cette définition se généralise à l'ellipse. Soient O_1 et O_2 deux points fixes dont la distance est $2c$. On appelle

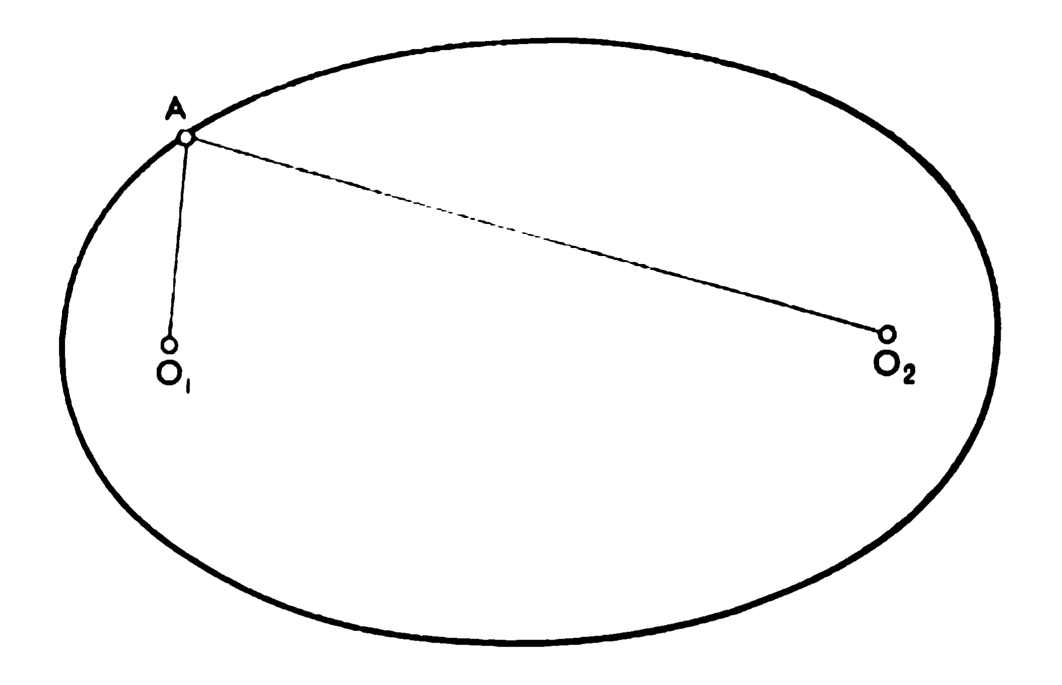


Fig. 4

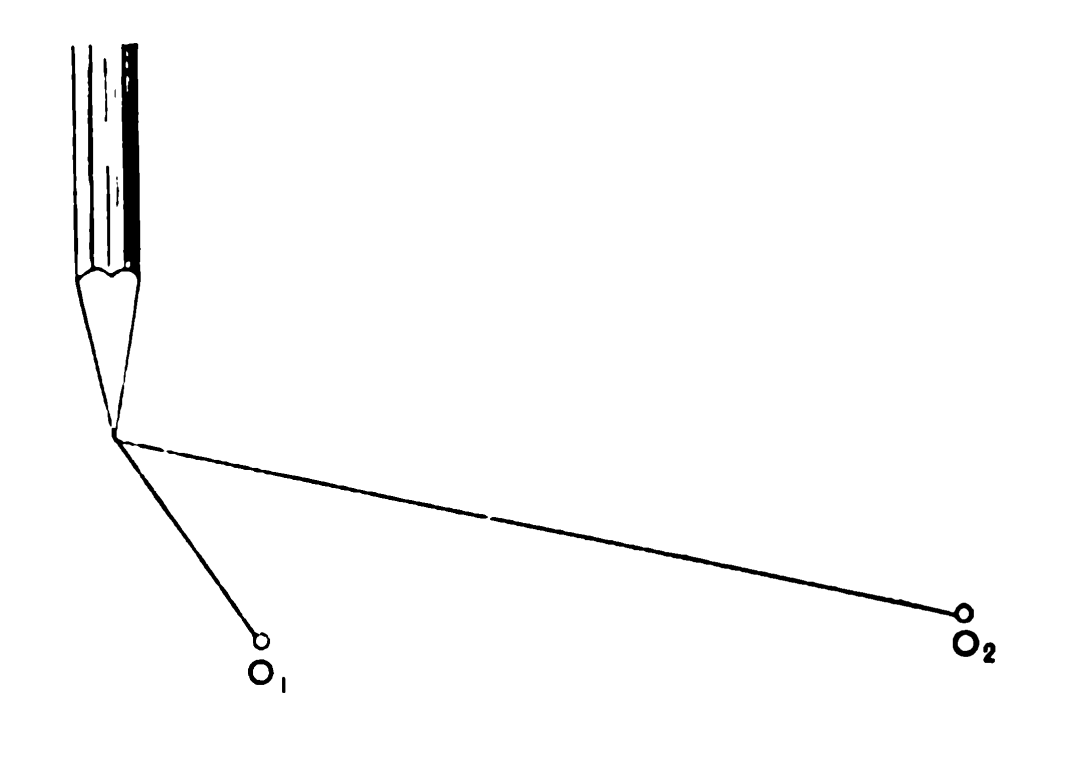


Fig. 5

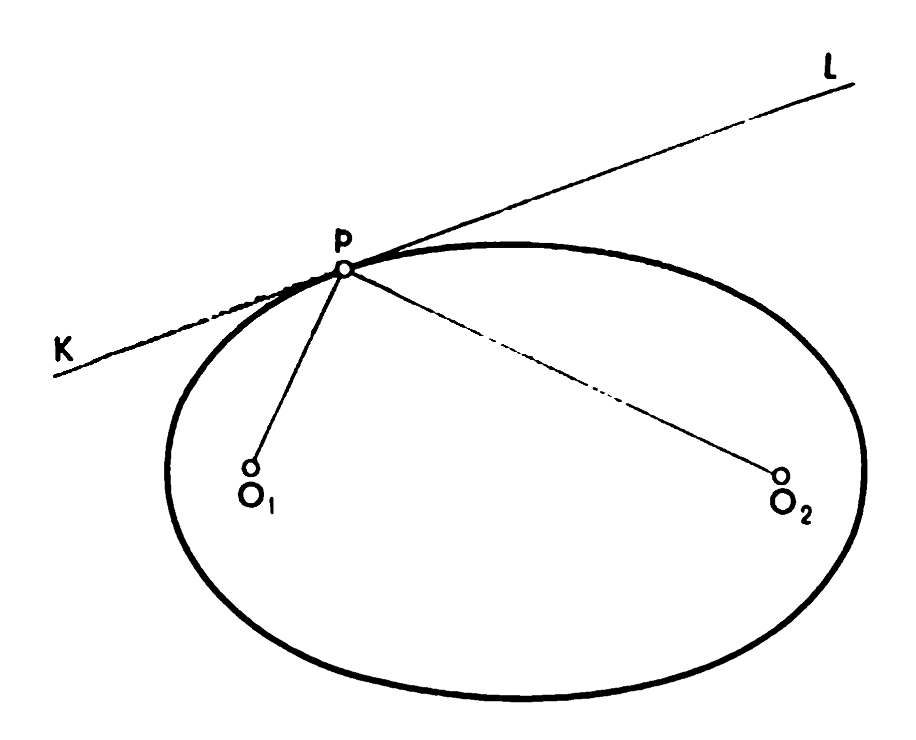


Fig. 6

ellipse (fig. 4) le lieu géométrique des points A dont la somme des distances aux points fixes O_1 et O_2 est égale à une constante $AO_1 + AO_2 = 2a$ supérieure à $2c$. Les points O_1 et O_2 sont les *foyers* de l'ellipse, $2a$ est la *longueur du grand axe*. Les segments AO_1 et AO_2 sont appelés *rayons vecteurs* du point A. Le cercle est un cas particulier de l'ellipse, celui précisément où les foyers O_1 et O_2 sont confondus. Les rayons vecteurs d'un point quelconque sont alors confondus et égaux chacun au rayon du cercle.

On obtient grossièrement le tracé d'une ellipse sur une feuille de papier en appuyant la pointe d'un crayon qui tend un fil fixé à deux points O_1 et O_2, tels que la distance à ces points soit inférieure à la longueur du fil (fig. 5). Pour plus de détails sur les propriétés de l'ellipse on renvoie le lecteur au cours de A. Markouchévitch : « Courbes remarquables », Editions Mir, 1973 qui fait partie du premier opuscule de cette série en français.

Comme pour le cercle on dira qu'une droite est tangente à une ellipse si elle ne possède qu'un point commun avec celle-ci (fig. 6).

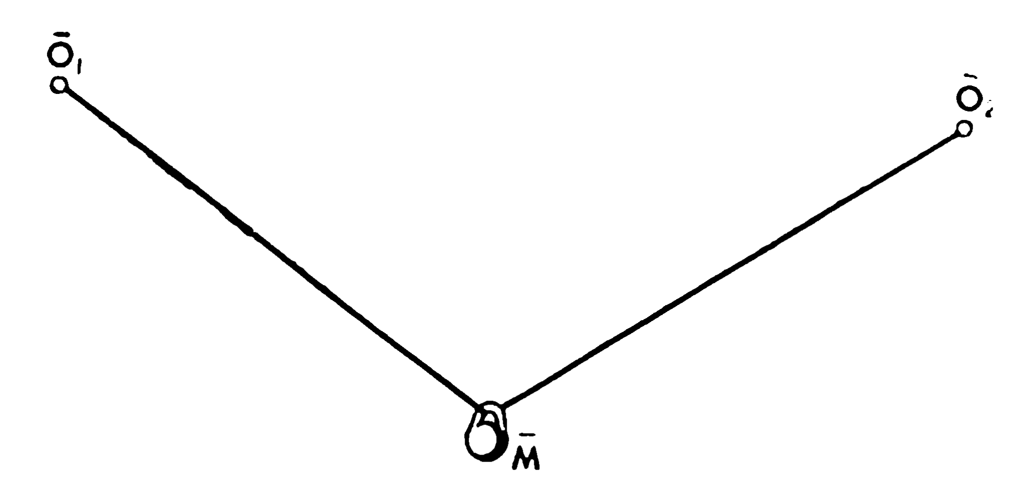


Fig. 7

Signalons l'existence du **théorème de la tangente à une ellipse**: *La tangente à une ellipse forme des angles égaux avec les rayons vecteurs du point de contact*[1]) (sur la fig. 6 les angles KPO_1 et LPO_2 sont donc égaux).

Cas particulier: le théorème de la tangente à un cercle. En effet, les rayons vecteurs d'un point du cercle étant confondus, ce théorème s'énonce ainsi: la tangente forme des angles adjacents égaux avec le rayon aboutissant au point de contact; or cela implique la perpendicularité de la tangente et du rayon[2]).

Nous allons maintenant démontrer le théorème de la tangente à une ellipse d'une façon analogue à celle du théorème de la tangente à un cercle. Nous recommandons préalablement au lecteur de s'exercer à imaginer une démonstration « mécanique ».

Comme dans le paragraphe précédent, nous allons remplacer la démonstration par une expérience. Soient deux

---

[1]) On peut énoncer ce théorème comme suit: la tangente en un point de l'ellipse est bissectrice extérieure de l'angle des rayons vecteurs de ce point. (*N.d.T.*)

[2]) On établit la perpendicularité de la tangente et du rayon en assimilant la première à la bissectrice extérieure de l'angle des rayons vecteurs du point de contact, et le second à la bissectrice intérieure du même angle (en effet, dans le cas du cercle les rayons vecteurs se confondent avec la bissectrice intérieure de l'angle qu'ils forment).

points fixes $\overline{O}_1$ et $\overline{O}_2$ situés dans un même plan vertical à égale hauteur. Prenons un fil de longueur supérieure à la distance des points $\overline{O}_1$ et $\overline{O}_2$ et fixons-le à ces points. Passons ce fil dans l'anneau d'un poids P comme l'indique la fig. 7. Si nous lâchons le poids P, il glissera le long du fil et finira par s'immobiliser en un point $\overline{M}$. Les deux propositions suivantes sont évidentes:

($\overline{\mathrm{C}}_1$) *Le plan* $\overline{O}_1\overline{O}_2\overline{M}$ *est vertical* (i.e. il contient une droite verticale).

($\overline{\mathrm{C}}_2$) *L'horizontale passant par le point* $\overline{M}$ *est bissectrice extérieure de l'angle formé par les segments* $\overline{MO}_1$ *et* $\overline{MO}_2$.

Remarquons que les propositions ($\overline{C}_1$) et ($\overline{C}_2$) pourraient être déduites de l'unicité de la position d'équilibre du poids P. Pour cela traçons la médiatrice du segment $\overline{O}_1\overline{O}_2$ et faisons pivoter le fil et le poids P de 180 autour de cette médiatrice. L'extrémité $\overline{O}_1$ du fil vient en $\overline{O}_2$ et l'extrémité $\overline{O}_2$ en $\overline{O}_1$. D'après la proposition (S) (voir paragraphe précédent) nous obtenons une position d'équilibre et, en vertu de l'unicité de celle-ci, le fil (et avec lui le triangle $\overline{M}\overline{O}_1\overline{O}_2$) occupe la même position qu'avant. Du moment que le triangle occupe la même position après avoir pivoté de 180°, cela veut dire que l'axe de rotation est situé dans le plan engendré par le triangle. Le plan $\overline{M}\overline{O}_1\overline{O}_2$ contenant une verticale est donc lui-même vertical. Par ailleurs, l'horizontale tracée à partir du point $\overline{M}$ dans le plan $\overline{M}\overline{O}_1\overline{O}_2$ se confond avec elle-même après pivotement et l'angle qu'elle forme avec le segment $\overline{MO}_1$ coïncide avec celui qu'elle forme avec le segment $\overline{MO}_2$. Elle est donc bissectrice extérieure de l'angle $\overline{O}_1\overline{M}\overline{O}_2$.

Supposons maintenant que les points O_1 et O_2 ne sont plus situés à une même hauteur. Passons un fil dans l'anse d'un poids P et fixons les extrémités de ce fil à O_1 et O_2 (fig. 8). Supposons que le poids P soit au repos en M. Remarquons que si le poids P s'est immobilisé en M et si nous fixons le fil en l'un quelconque de ses points, la position de repos reste la même. (En effet, si nous fixons le fil en $\overline{O}_2$ comme l'indique la fig. 8, la tension du brin de fil $\overline{O}_2O_2$

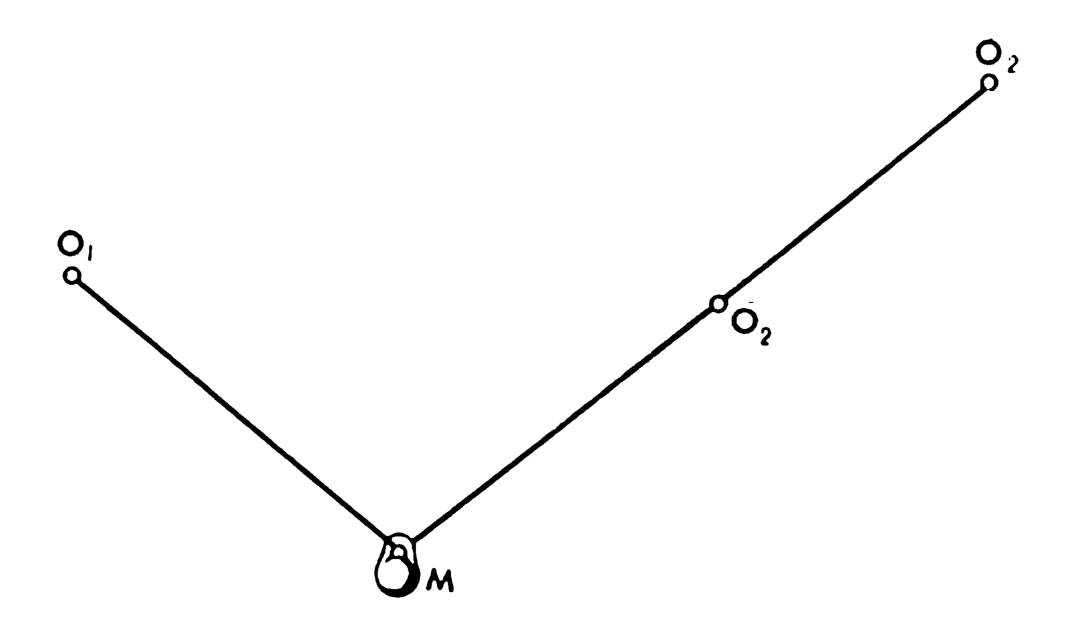


Fig. 8

est remplacée par la réaction du support au point $\bar{O}_2$.) Ceci nous permet de généraliser les conclusions faites dans le cas où les points $\bar{O}_1$ et $\bar{O}_2$ étaient situés à une même hauteur. En effet, il suffit d'accrocher le fil en un point $\bar{O}_2$ situé à une même hauteur que le point O_1 (fig. 8[1])); la position d'équilibre n'en sera pas modifiée. En posant $\bar{O}_1 = O_1$, nous nous plaçons dans le cas examiné précédemment de deux points situés à une même hauteur. Donc les propositions $(\bar{C}_1)$ et $(\bar{C}_2)$ entraînent aussitôt la proposition $\bar{C}_1$ (qui est évidente) et la proposition $(\bar{C}_2)$:

(C_1) *Le plan O_1O_2M est vertical.*

(C_2) *Dans le plan MO_1O_2, l'horizontale qui passe par le point M est bissectrice extérieure de l'angle formé par les segments MO_1 et MO_2.*

Enfin les deux propositions suivantes sont évidentes:

(A) *Lorsqu'il est en équilibre, un poids suspendu à un fil occupe sa position la plus basse* (si l'on utilise le principe du

---

[1]) Sur la fig. 8 le point O_2 est plus élevé que le point O_1; le cas contraire se traite d'une façon analogue.

minimum de l'énergie potentielle, cette proposition découle de l'unicité de l'état d'équilibre).

(B) *Lorsque l'équilibre est réalisé, le fil est tendu* (i.e. les brins de fil MO_1 et MO_2 sont rectilignes et la somme des distances du point M aux points O_1 et O_2 est égale à la longueur du fil).

Revenons maintenant à la démonstration de notre théorème de la tangente. Soit une ellipse de foyers O_1 et O_2 (voir fig. 6). On désigne par l la longueur du grand axe. On trace une droite KL tangente en P à cette ellipse. On demande de prouver qu'elle est bissectrice extérieure de l'angle $\widehat{O_1PO_2}$. Pour les besoins de la démonstration tournons le dessin de façon que le plan qui le contient soit vertical, la droite KL horizontale, et l'ellipse située au-dessus de KL (le point K sera alors le point le plus bas de cette ellipse). Soit un fil de longueur l. Passons-le dans l'anse d'un poids et fixons ensuite ses extrémités en O_1 et O_2. Si nous lâchons ce poids, il glissera le long du fil et finira par s'immobiliser en un point M. En vertu de (C_1) ce point sera situé dans le plan du dessin. En vertu de (B), il sera porté par l'ellipse. En vertu de (A) il sera confondu avec P. Donc, le fil se tendra suivant les segments PO_1 et PO_2. En vertu de (C_2) la droite KL est bissectrice extérieure de l'angle formé par ces segments.

Le lecteur trouvera une autre démonstration fondée sur des notions de mécanique dans le § 11 du livre de L. Lusternick « Les lignes courtes ».

On construit à la règle et au compas la tangente à une ellipse en un point P de celle-ci, en utilisant la propriété de la tangente d'être bissectrice extérieure de l'angle des rayons vecteurs du point P (voir fig. 9).

Le théorème de la tangente à l'ellipse se connaît une intéressante interprétation optique: tout rayon lumineux passant par l'un des foyers est réfléchi par la paroi de l'ellipse dans l'autre foyer (fig. 10).

En effet, les angles formés avec l'ellipse par le rayon incident et le rayon réfléchi sont égaux. L'angle formé par une droite et une

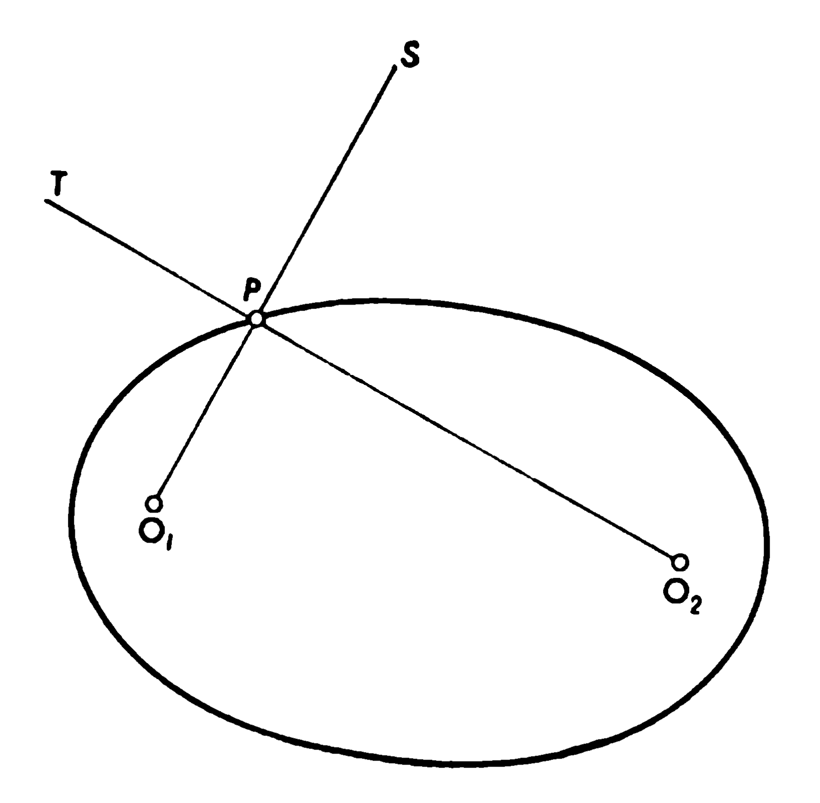


Fig. 9

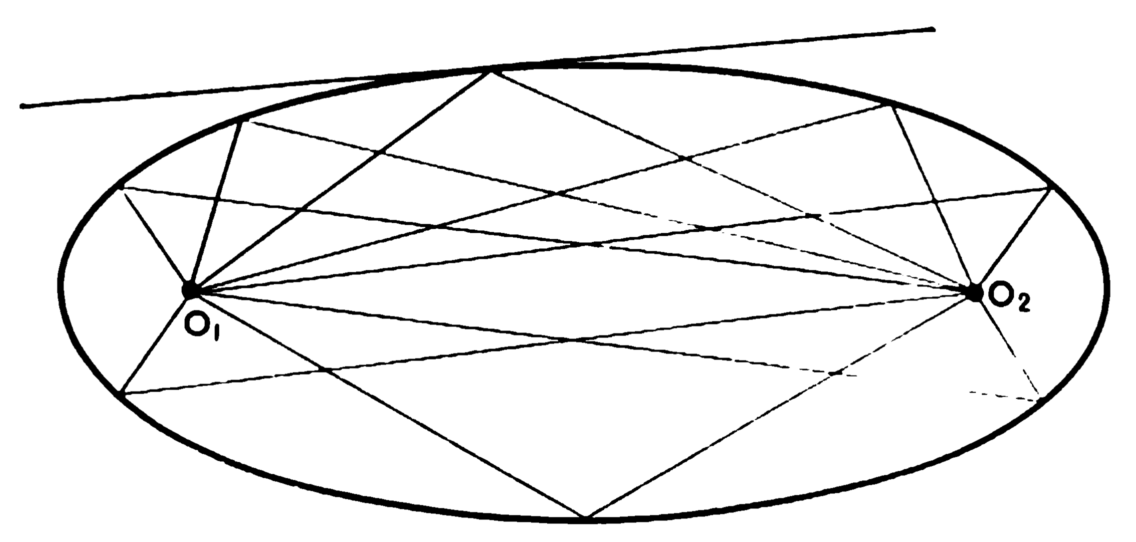


Fig. 10

courbe (dans le cas présent par le rayon lumineux et l'ellipse) est l'angle formé par cette droite avec la tangente à l'ellipse au point de contact de cette droite avec l'ellipse. Puisque le rayon incident est un rayon vecteur, en vertu du théorème démontré, le rayon réfléchi sera l'autre rayon vecteur du point de contact avec l'ellipse.

§ 3 PROBLÈME DE LA TANGENTE À UNE PARABOLE ET À UNE HYPERBOLE

La *parabole* (fig. 11) est le lieu géométrique des points P situés à égale distance d'un point fixe F appelé foyer et d'une droite d appelée *directrice.* (Pour plus de détails sur la parabole on renvoie le lecteur à l'ouvrage déjà mentionné de Markouchévitch). La parabole divise le plan en deux régions, l'une contenant la directrice, l'autre (hachurée sur la fig. 12) le foyer. On appelle *tangente* à la parabole toute droite possédant un point commun avec cette parabole et contenue dans l'une des deux régions. Sur la fig. 13 la droite p est une tangente contrairement à la droite q bien que cette dernière ne possède qu'un point commun avec la parabole. On appelle rayons vecteurs d'un point P de la parabole le rayon PF et le rayon PD où D est la projection orthogonale de P sur la directrice (d).

Nous avons le théorème suivant relatif à la tangente à une parabole: *la tangente à une parabole en un point P* (fig. 14) *de celle-ci est bissectrice intérieure de l'angle formé par les rayons vecteurs PF et PD du point P.*

On peut «démontrer» ce théorème en faisant appel comme dans les paragraphes précédents à des notions de mécanique. Esquissons cette « démonstration ».

A une distance quelconque l de la droite (d) traçons une droite (d') parallèle à (d) et telle que le point P soit compris entre (d) et (d'). Amenons le dessin dans un plan vertical de façon que la tangente (p) soit horizontale et la parabole située au-dessus de (p). Nous assimilerons la droite (d') à une tige fine le long de laquelle glisse un anneau (fig. 15) attaché à un fil inextensible de longueur l dont l'autre

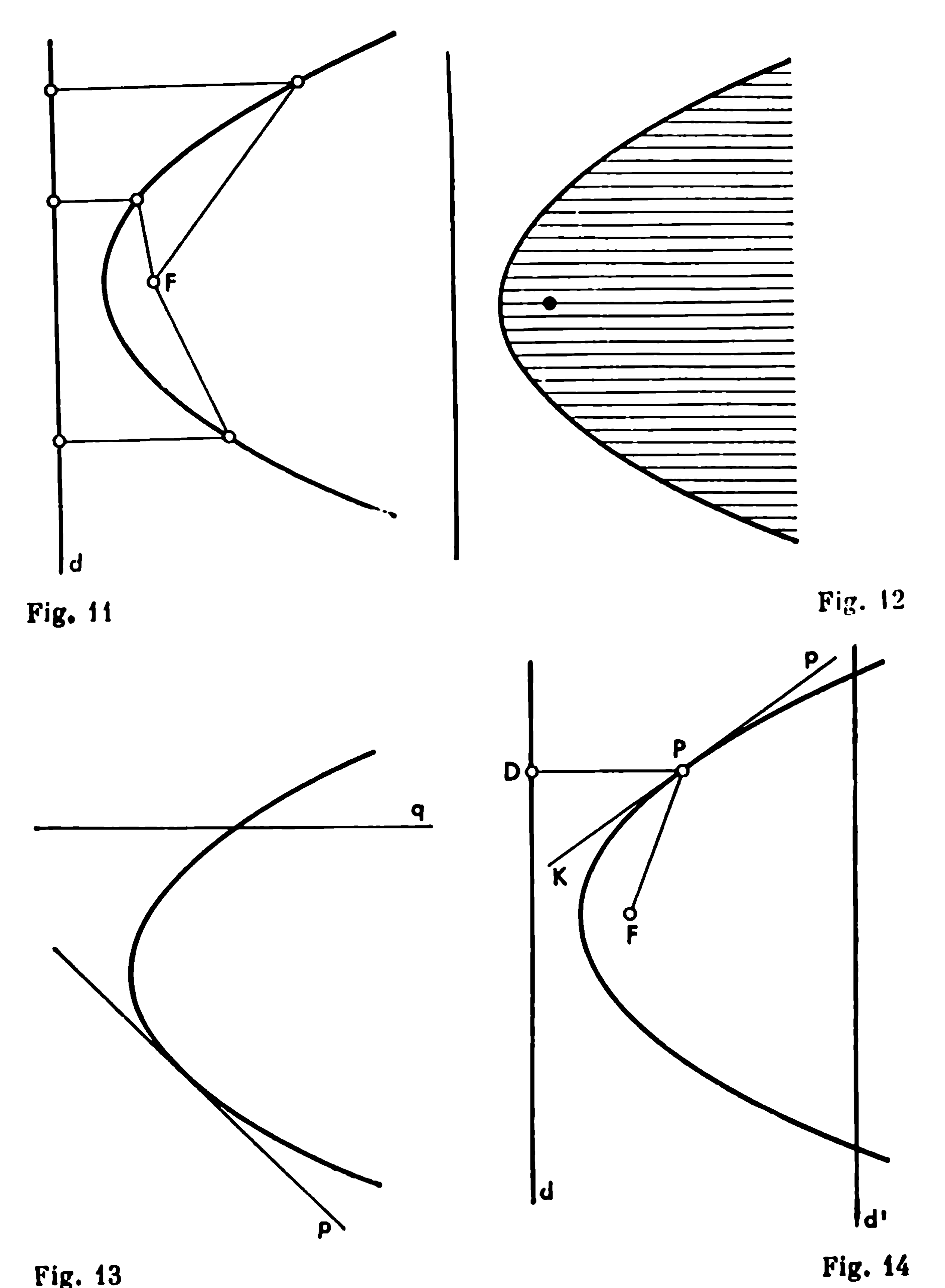


Fig. 11

Fig. 12

Fig. 13

Fig. 14

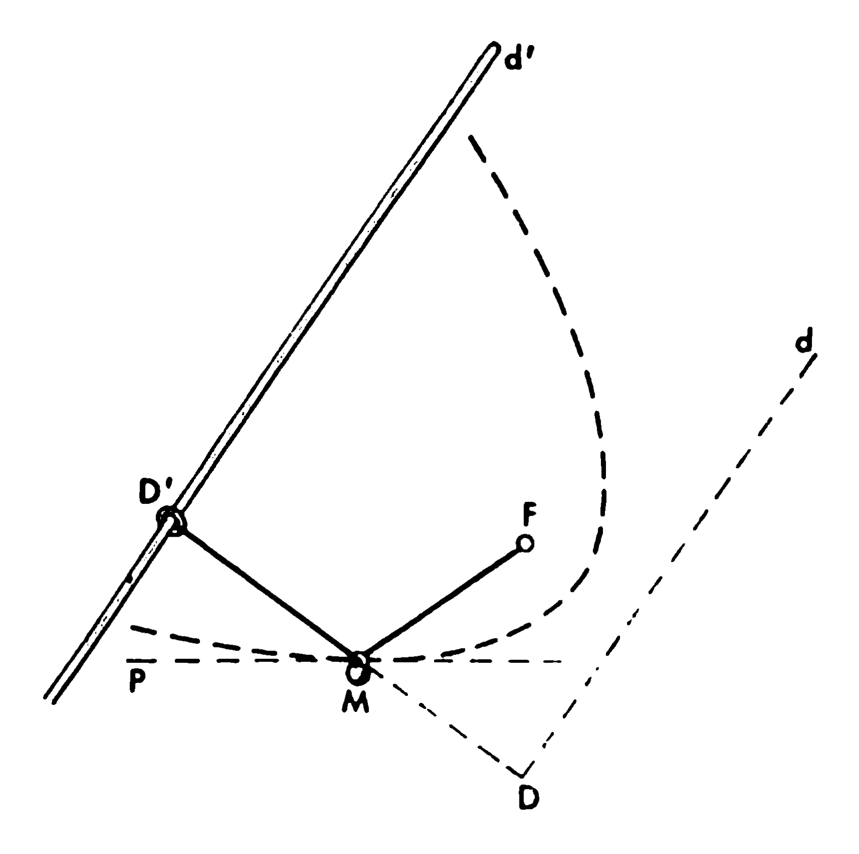


Fig. 15

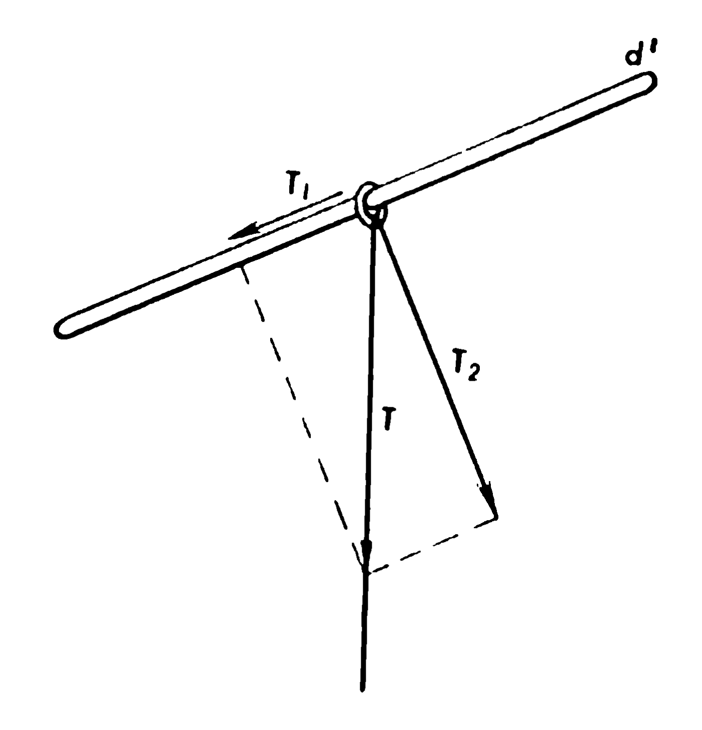


Fig. 16

extrémité est fixée en F. Passons ce fil dans l'anse d'un poids. Supposons que le poids se trouve au repos en M ; l'anneau se trouvera alors en D'. Le fil étant tendu, les brins MF et MD' sont rectilignes. Remarquons tout d'abord que

$$MD' \perp (d'). \tag{1}$$

(S'il en était autrement, ainsi que le montre la fig. 16, la composante T_1 de la force T aurait été différente de zéro et aurait mis l'anneau en mouvement.) D'autre part, en reprenant les raisonnements du paragraphe précédent, on établit d'abord que la droite (p) forme des angles égaux avec les segments MD' et MF et ensuite que

$$MD' + MF = l. \tag{2}$$

Menons $MD \perp (d)$. La distance entre d et d' étant égale à l, de (1) et (2) il résulte que

$$MD = MF. \tag{3}$$

Ce qui traduit le fait que le point M est situé sur la parabole et comme il occupe sa position la plus basse et que P est le point le plus bas de la parabole, M est confondu avec P. Finalement, les angles formés par les segments PF et PD avec la droite sont égaux, i.e. (p) est bissectrice intérieure de l'angle des rayons vecteurs MF et MD du point M.

Dans le livre de L. Lusternick « Les lignes courtes » est développée une autre démonstration « mécanique » de ce théorème.

La propriété de la tangente d'être bissectrice intérieure de l'angle des rayons vecteurs permet de la construire à l'aide du compas et de la règle.

On peut donner l'interprétation optique suivante au théorème démontré : si l'on place une source lumineuse dans le foyer d'une parabole, les rayons réfléchis par les parois de celle-ci se propagent suivant une direction perpendiculaire à la directrice (fig. 17). Cette propriété de la parabole est utilisée dans les projecteurs, réflecteurs et autres dispositifs propres à réfléchir les rayons (lumineux ou thermi-

ques) dans une direction donnée : la surface réfléchissante de ces dispositifs est la surface engendrée par la rotation d'une parabole autour de son axe.

L'hyperbole est le lieu géométrique des points dont la valeur absolue de la différence des distances à deux points fixes F_1 et F_2, appelés *foyers*, est constante. (Pour plus de détails on renvoie le lecteur au cours de A. Markouchévitch « Courbes remarquables ».) On voit donc sur la fig. 18 que l'hyperbole est constituée de deux branches, qui, comme dans le cas de la parabole, divisent le plan en deux régions. Toute droite entièrement contenue dans l'une des deux régions du plan et possédant un point commun avec la branche correspondante est appelée *tangente à la branche de l'hyperbole.* Sur la fig. 19 la droite KL est tangente en P à la branche de gauche. On démontre qu'une tangente à une branche ne coupe pas l'autre branche. C'est pourquoi la tangente à une branche est appelée tout simplement *tangente à l'hyperbole.*

On appelle *rayons vecteurs* d'un point les segments qui le joignent aux foyers de l'hyperbole.

Nous avons le théorème suivant relatif à la tangente à une hyperbole : la tangente à l'hyperbole *en un point quelconque de celle-ci est bissectrice intérieure de l'angle des rayons vecteurs de ce point.*

Nous laissons au lecteur le soin de démontrer ce théorème ; pour notre part nous nous contenterons simplement de lui suggérer une méthode « mécanique » de démonstration. D'après la définition de l'hyperbole, la valeur absolue de la différence des longueurs des rayons vecteurs est constante ; désignons cette différence par a. Supposons que le point de contact est situé sur la branche droite (voir fig. 20). Décrivons un cercle de centre F_1 et de rayon $r > a$ et tel que le point P soit situé à l'intérieur de ce cercle.

Assimilons ce cercle à un fil de fer suivant lequel glisse un anneau. Passons un fil de longueur $r - a$ dans l'œillet d'un poids et fixons l'une de ses extrémités au foyer F_2 et

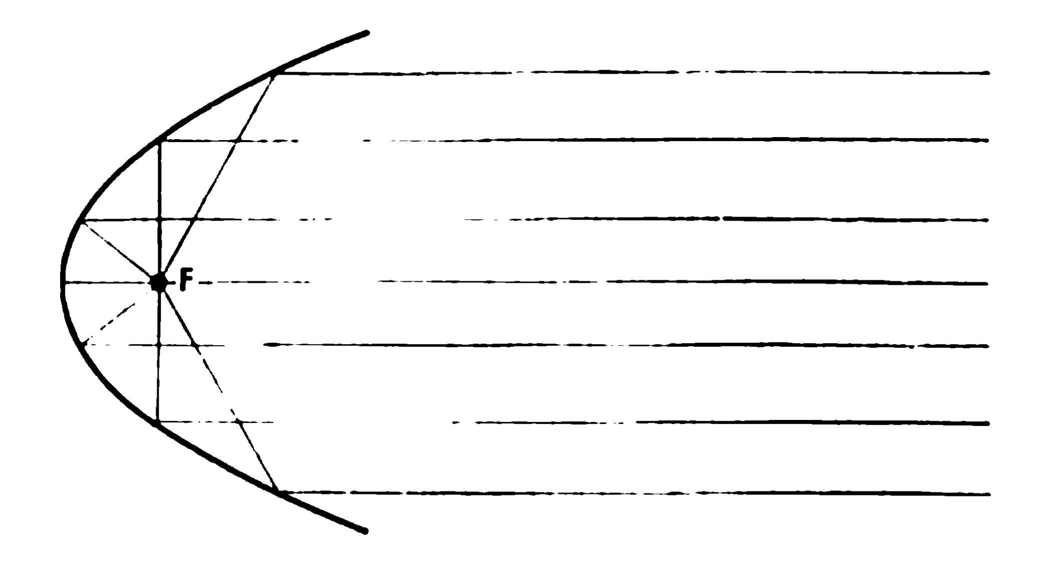


Fig. 17

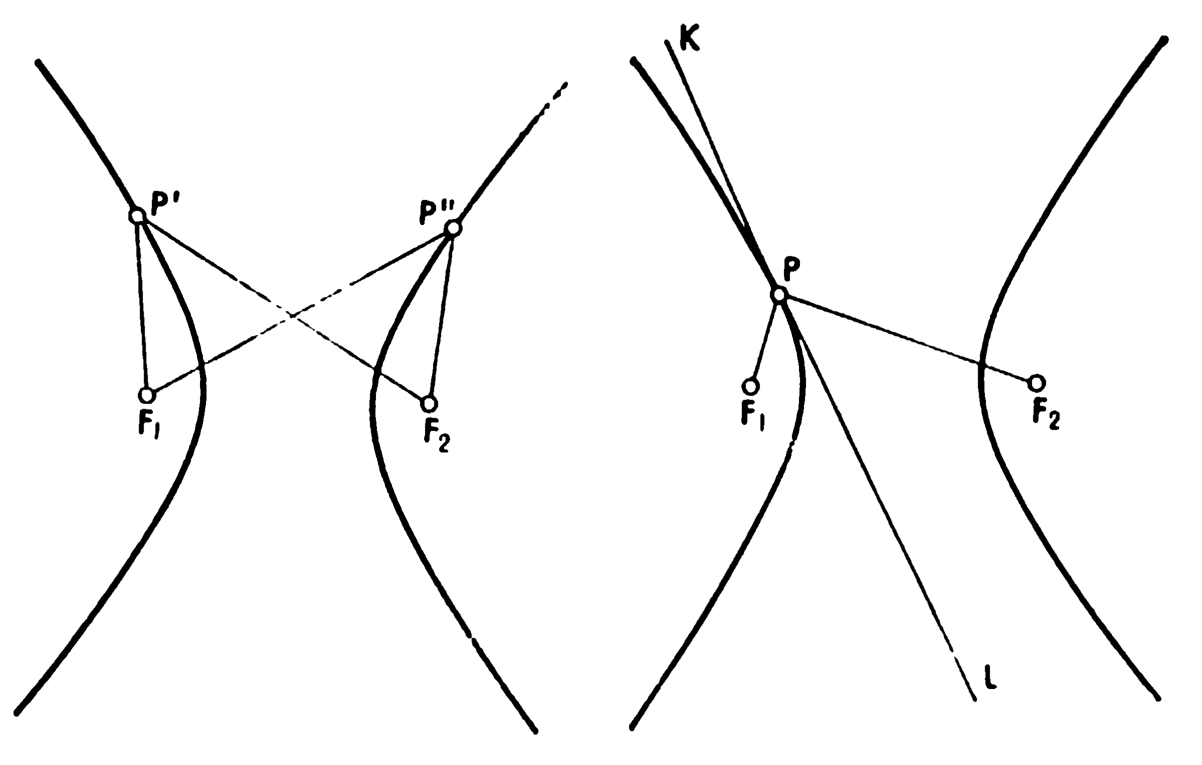


Fig. 18

Fig. 19

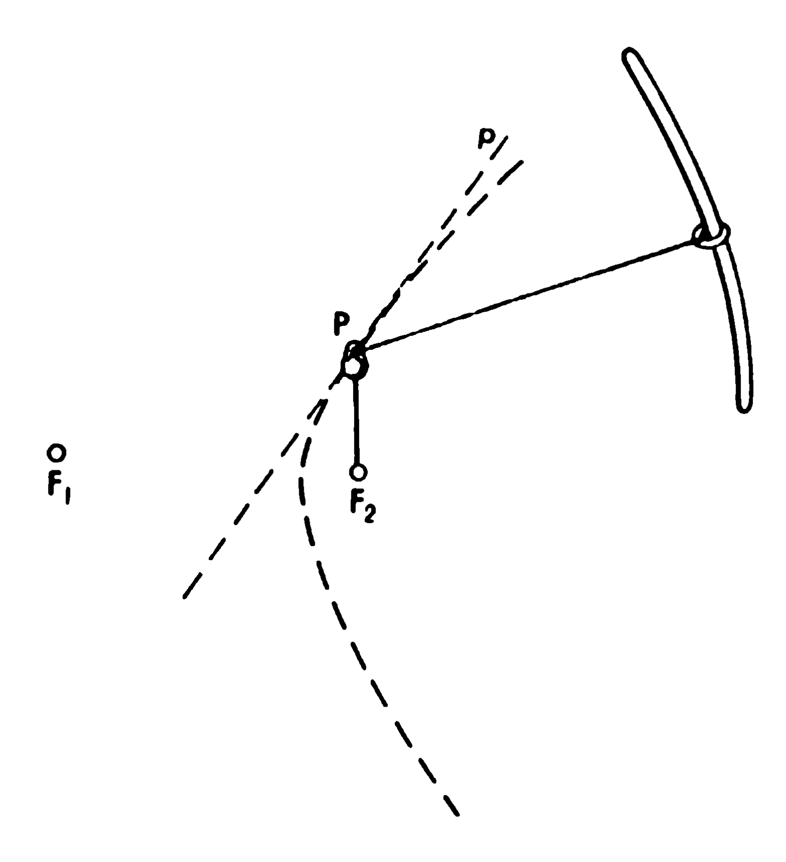


Fig. 20

l'autre extrémité à l'anneau. Si maintenant on tourne le dessin de manière que la droite (p) (la tangente) soit horizontale et la branche correspondante de l'hyperbole située au-dessus de (p), le poids glissera le long du fil et s'immobilisera en P. Cet artifice nous permet de démontrer ce théorème en reprenant les raisonnements précédents.

Si le point de contact se trouve sur la branche gauche, on refait la même construction en prenant cette fois-ci pour centre du cercle le foyer F_2. On peut toutefois se servir du premier cercle à condition que la longueur du fil soit égale à $r + a$.

Comme dans le cas de l'ellipse et la parabole, le théorème de la tangente nous fournit en même temps une méthode de sa construction à la règle et au compas.

§ 4 PRINCIPE DU MINIMUM DE L'ÉNERGIE POTENTIELLE

Toute masse élevée à une certaine hauteur est susceptible en tombant de fournir un travail, i.e. elle possède de l'énergie potentielle. Du cours élémentaire de physique on sait qu'un corps de poids q placé à une hauteur h a une énergie potentielle qh. Nous voyons que pour un corps donné plus h est petit plus l'énergie potentielle est faible. L'énergie potentielle a toujours tendance à diminuer, c'est pourquoi tout corps pesant tend à occuper sa position la plus basse. Si ce corps est attaché à un fil, à sa position la plus basse son énergie potentielle sera minimale. C'est pourquoi la proposition (A), capitale dans la démonstration des théorèmes relatifs aux tangentes et selon laquelle à l'état de repos tout corps pesant occupe sa position la plus basse, est équivalente à la proposition suivante:

(U) *Si un solide se trouve en état d'équilibre, son énergie potentielle prend sa valeur minimale.*

La proposition (U) (et donc la proposition (A)) est à son tour une conséquence de la proposition (E) relative à l'unicité de l'état d'équilibre et de la proposition suivante:

(D') *Si l'énergie potentielle d'un solide prend sa valeur minimale, celui-ci se trouve en état d'équilibre.*

Pour déduire (U) il suffit de remarquer que si à l'état d'équilibre l'énergie potentielle n'était pas minimale, en vertu de (D') il existerait un autre état d'équilibre qui correspondrait à la valeur minimale de l'énergie potentielle; or, cela contredit la proposition (E).

La proposition (D') est un cas particulier d'un principe général de la mécanique appelé principe du minimum de l'énergie potentielle ou encore principe de Dirichlet. Le principe de Dirichlet s'énonce comme suit:

Si l'énergie potentielle d'un système est minimale, celui-ci est en état d'équilibre.

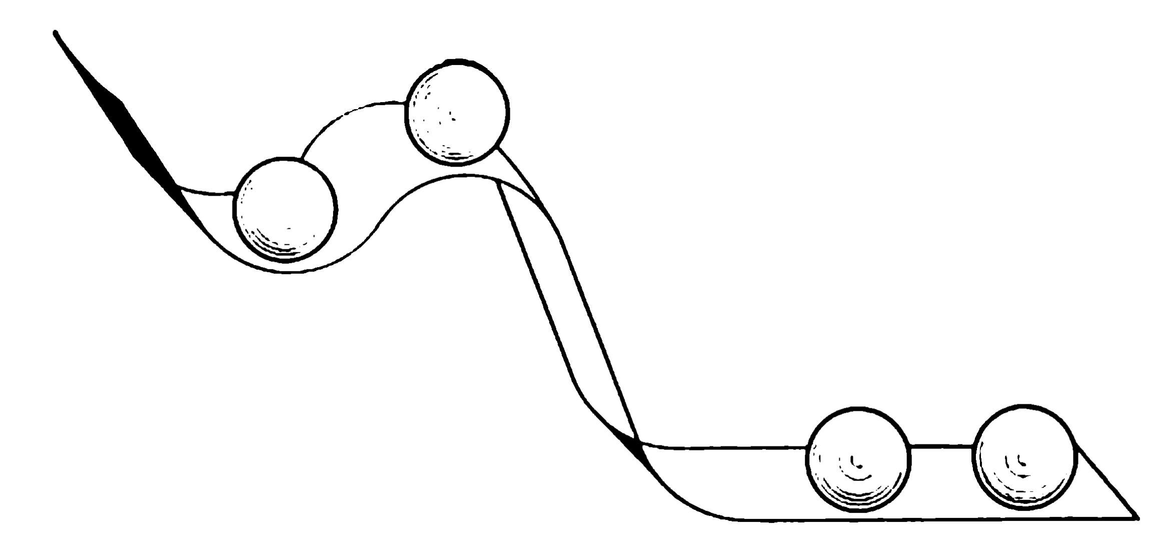


Fig. 21

Si cet état d'équilibre est unique, le principe de Dirichlet admet une conséquence importante :

L'énergie potentielle du système en état d'équilibre est minimale.

Cette conséquence se déduit comme la proposition (U).

Dans la suite nous ne considérerons que les cas où l'état d'équilibre est unique (d'autres cas sont éventuellement possibles ; ainsi, sur la fig. 21 on voit que la boule possède quatre états d'équilibre).

Il n'y a certes plus lieu de recourir au principe du minimum de l'énergie potentielle pour résoudre les problèmes relatifs aux tangentes ; le fait que le solide pesant occupe sa position la plus basse est évident en soi. Cependant dans nombre de cas lorsque nous avons affaire non plus à un seul mais à plusieurs solides liés entre eux, il est incorrect d'affirmer qu'en état d'équilibre chacun d'eux occupe sa position la plus basse ; dans ce cas il importe de définir l'état de tous les corps, et pour cela il est souvent commode d'introduire la notion d'énergie potentielle.

Illustrons ce que nous venons de dire sur un exemple.

Dans le remarquable livre du mathématicien polonais

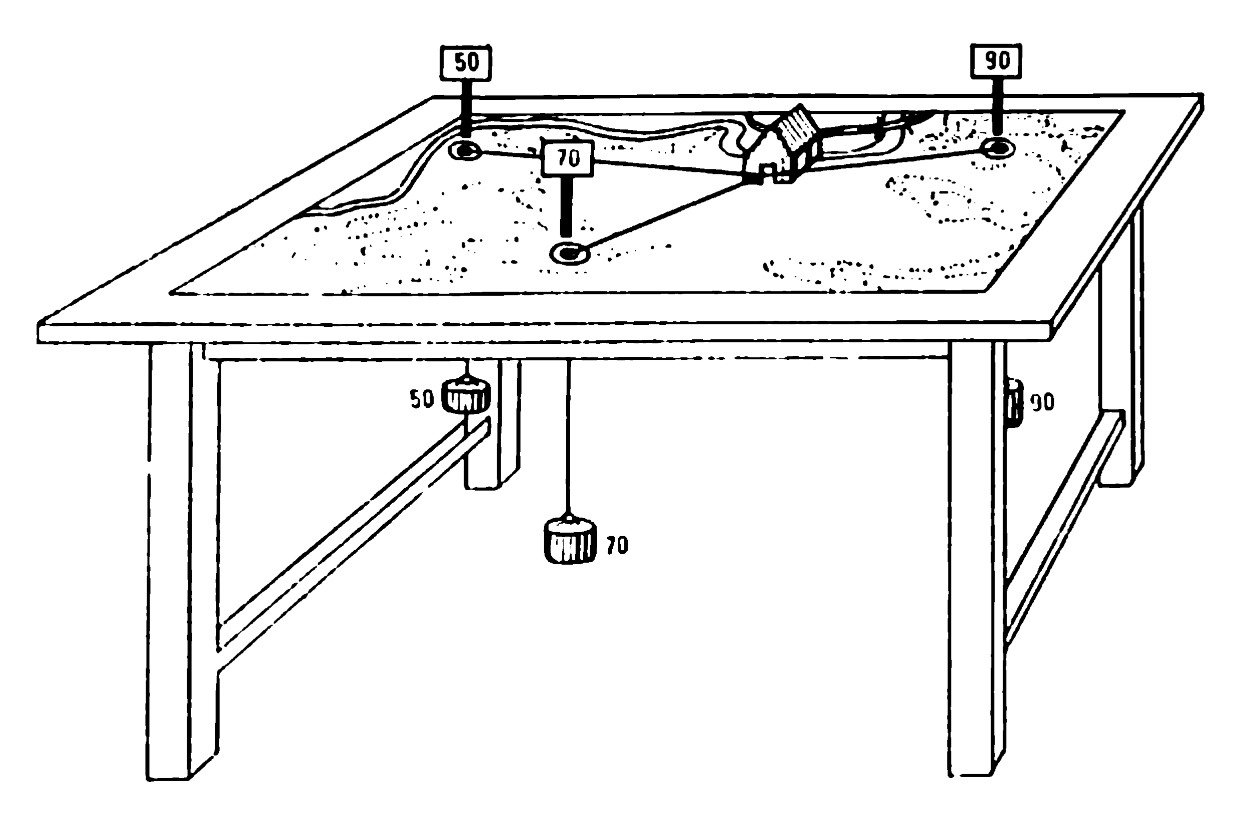


Fig. 22

G. Steinhauss « Kaléidoscope de mathématiques » (Gostekhizdat M.-L. 1949) figure le problème suivant (sujet 39).

« On se propose de construire une école appelée à desservir trois communes. 50 élèves habitent dans la première commune, 70 dans la seconde et 90 dans la troisième. Où doit-on construire cette école pour que la somme des temps dépensés par les élèves pour s'y rendre soit minimale?

« Pour résoudre ce problème il suffit d'étendre sur la table le plan de la région (fig. 22), de percer trois trous à l'emplacement respectif des communes, d'y passer trois fils dont les extrémités supérieures sont nouées ensemble et aux extrémités inférieures desquels on suspend des charges de poids respectivement proportionnel à 50, 70 et 90 (on prendra, par exemple, 500, 700 et 900 grammes). L'école devra être construite à l'endroit où s'immobilisera le nœud. On demande pourquoi ?»

Pour répondre à cette question nous allons calculer l'énergie potentielle du système formé par les trois charges. Si q_1 est le poids de la première charge, q_2 celui de la deuxième et q_3 celui de la troisième et si elles sont respectivement à une hauteur de h_1, h_2 et h_3, l'énergie potentielle du système qu'elles forment est égale à la somme des énergies potentielles de chacune d'elles[1]:

$$E = q_1h_1 + q_2h_2 + q_3h_3. \quad (1)$$

Supposons maintenant que r_1, r_2 et r_3 désignent respectivement les distances du nœud à la première, la seconde et la troisième commune; l_1, l_2 et l_3 les longueurs des premier, second et troisième fils et h la hauteur de la table. Il est évident que quelle que soit la position occupée par le nœud et quelles que soient les hauteurs h_1, h_2 et h_3 des charges, on a les relations

$$r_1 + (h - h_1) = l_1, \quad r_2 + (h - h_2) = l_2, \quad r_3 + (h - h_3) = l_3$$

ou

$$h_1 = r_1 + h - l_1, \quad h_2 = r_2 + h - l_2, \quad h_3 = r_3 + h - l_3.$$

La relation (1) peut s'écrire:

$$E = q_1r_1 + q_2r_2 + q_3r_3 + C,$$

où

$$C = (q_1 + q_2 + q_3)h - q_1l_1 - q_2l_2 - q_3l_3$$

est une quantité constante ne dépendant pas de la position des charges.

Si E prend sa valeur minimale, en vertu du principe du minimum de l'énergie potentielle le système se trouve en

---

[1]) En effet, l'énergie potentielle se mesure par le travail fourni. Si nous coupons les fils, en tombant, chaque charge effectuera un travail égal à son énergie potentielle, et le système tout entier, un travail égal à la somme des travaux effectués séparément par chaque élément de ce système.

état d'équilibre. D'où il résulte — sous la condition que l'état d'équilibre soit unique — que la quantité E est minimale à l'état d'équilibre. Or, si E est minimale, la quantité

$$T = q_1r_1 + q_2r_2 + q_3r_3 = E - C$$

est aussi minimale. Or, T est précisément la somme des temps mis par les enfants pour se rendre de leurs communes à l'école si celle-ci est située à l'emplacement du nœud. Donc lorsque le système est en état d'équilibre, le nœud occupe une position telle que la somme des temps mis pour l'atteindre est minimale.

§ 5 POINTS MATÉRIELS ET CENTRE DE GRAVITÉ

Jusqu'à maintenant nulle part il n'a été question des dimensions des solides suspendus aux divers fils. Dans nos raisonnements d'un côté nous avons implicitement admis que les solides étaient de faibles dimensions (si faibles qu'on pouvait les assimiler à des points) et de l'autre, nous les avons dotés de poids, donc de masses.

Ceci nous amène à considérer l'une des plus importantes notions de mécanique, la notion de point matériel. Par point matériel on entend un solide de dimensions négligeables. On peut se représenter un point matériel comme un point géométrique affecté d'un nombre (appelé « masse »). Remarquons que du moment que le point matériel possède une masse, il possède également un poids.

Si nous avons un système de points matériels, la pesanteur exerce sur chacun d'eux une force proportionnelle à leurs masses. La résultante de toutes ces forces parallèles est appliquée en un point appelé centre de gravité du système de points matériels. La position du centre de gravité est fonction de la position des points matériels du système et de leurs masses.

En mécanique, tout solide est considéré comme étant constitué d'un grand nombre de points matériels. Ceux-ci étant en règle générale de masses différentes, la masse est inégalement répartie à l'intérieur du solide: certaines parties du solide possédant une masse supérieure à celle des autres. Nous irons même jusqu'à admettre que certaines parties du solide peuvent ne posséder aucune masse c'est-à-dire sont « impondérables »[1]). Le centre de gravité d'un système de points matériels formant un solide est par définition le centre de gravité de ce solide. Si la masse de ce dernier est concentrée en un nombre fini de ses points (ces cas se présenteront dans la suite et notamment dans une tige de masse négligeable à laquelle sont suspendus des poids ponctuels), le centre de gravité de ce solide est de toute évidence confondu avec le centre de gravité du système formé par ces points.

Beaucoup de propriétés du solide dépendent exclusivement de la position de son centre de gravité. On sait du cours élémentaire de physique que l'état d'équilibre d'un solide dépend des positions respectives de son centre de gravité et de l'ensemble de ses points fixes. Nous avons notamment les propositions suivantes:

Proposition I. *Si un solide suspendu par l'un de ses points se trouve à l'état de repos, ce point et le centre de gravité du solide sont situés sur une même verticale* (fig. 23).

Proposition II. *Si un solide repose sur un plan par l'une de ses faces ou l'une de ses arêtes, la verticale qui passe par le centre de gravité coupe cette face ou cette arête* (fig.24).

---

[1]) Plus précisément, par solide impondérable (ou solide de « masse nulle ») nous entendons un solide de masse si petite qu'on peut la négliger. Le lecteur ne manquera pas de faire un parallèle entre la notion de point matériel et celle de solide impondérable. La négligence de la masse du solide joue en mécanique un rôle aussi important que la négligence des dimensions de ce solide. Ainsi, dans nos précédents raisonnements nous avons négligé les poids (et non pas les dimensions!) des fils.

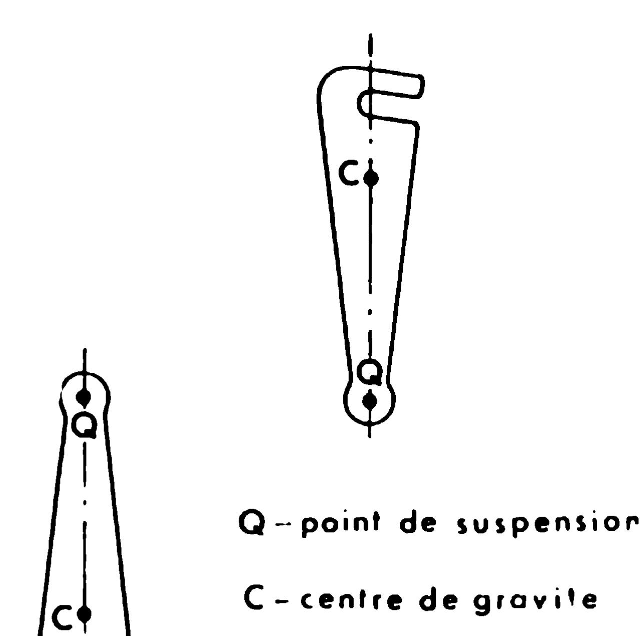


Fig. 23

De ces propositions (qui, notons-le, peuvent servir à déterminer expérimentalement le centre de gravité d'un solide) il résulte en particulier: si on remplace un solide par un autre, de même forme et de même centre de gravité, l'état d'équilibre n'est pas modifié. Cependant, dans nombre de cas il est indispensable, outre la position du centre de gravité, de tenir compte de la masse totale du solide (i.e. la masse de tous les points matériels qui le composent). Si, par exemple, nous remplaçons une partie d'un solide par un autre solide ayant la même forme et le même centre de gravité qu'elle, il peut arriver que l'équilibre du premier solide soit troublé à cause de la variation de masse. La connaissance du centre de gravité et de la masse totale suffit pour résoudre de nombreux problèmes, du moins ceux auxquels nous serons confrontés dans la suite. Aussi introduirons-nous la définition suivante d'équivalence de deux systèmes de points matériels:

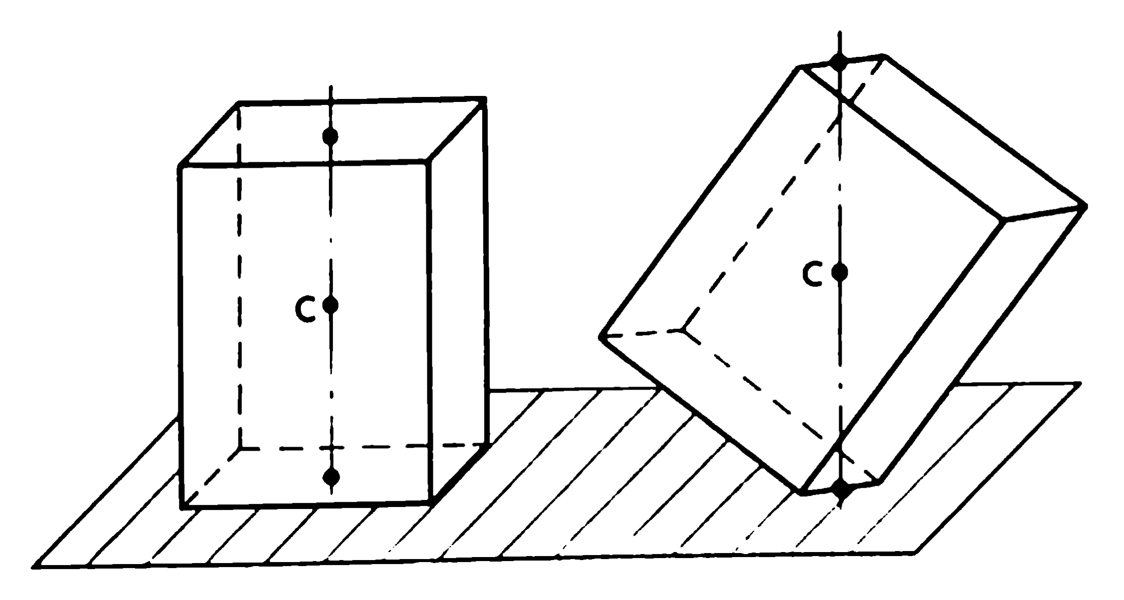


Fig. 24

Deux systèmes de points matériels sont équivalents premièrement, si le centre de gravité du premier système est confondu avec celui du second, et, deuxièmement, si la somme des masses des points du premier système est égale à la somme des masses des points du second système.

La notion d'équivalence que nous venons de définir est remarquable en ce sens que lorsqu'on remplace une partie d'un système par un système équivalent, nous obtenons un nouveau système équivalent au système initial. Ce que l'on peut encore formuler d'une manière plus rigoureuse.

Soient $\bar{\alpha}$ un système de points matériels $\bar{M}_1, \ldots, \bar{M}_r$ de masses respectivement $\bar{m}_1, \ldots, \bar{m}_r$ et $\underline{\alpha}$ un système de points matériels $\underline{M}_1, \ldots, \underline{M}_s$ respectivement de masses $\underline{m}_1, \ldots, \underline{m}_s$. Groupons ces deux systèmes en un seul système α que nous définirons comme suit:

1° Tout point matériel du système $\bar{\alpha}$ distinct, par sa position, de tout point du système $\underline{\alpha}$ fera partie de α.

2° Tout point matériel du système $\underline{\alpha}$ distinct, par sa position, de tout point du système $\bar{\alpha}$ fera partie de α.

3° Si un point matériel $\overline{M}_i$ du système $\overline{\alpha}$ est confondu avec un point quelconque $\underline{M}_j$ du système $\underline{\alpha}$, nous inclurons dans α un nouveau point matériel confondu avec $\overline{M}_i$ et $\underline{M}_j$, mais de masse égale à la somme $\overline{m}_i + \underline{m}_j$ des masses de ces deux points. Nous avons:

P r o p o s i t i o n III. *Si le système α est la réunion des systèmes $\overline{\alpha}$ et $\underline{\alpha}$ et que le système β soit la réunion des systèmes $\overline{\beta}$ et $\underline{\beta}$ et si les systèmes $\overline{\alpha}$ et $\overline{\beta}$ d'une part et $\underline{\alpha}$ et $\underline{\beta}$ d'autre part sont équivalents, alors les systèmes α et β sont également équivalents.*

D é m o n s t r a t i o n d e l a p r o p o s i t i o n III. Il est évident que les masses totales des systèmes α et β sont égales. Montrons que ces systèmes possèdent le même centre de gravité. Pour trouver la résultante des forces de pesanteur agissant sur les points matériels du système α, on peut procéder comme suit: on détermine d'abord la résultante $\overline{P}$ des forces agissant sur $\overline{\alpha}$, puis la résultante $\underline{P}$ des forces agissant sur $\underline{\alpha}$ et, enfin, la résultante P de ces deux résultantes. D'une manière analogue, on détermine la résultante Q des forces $\overline{Q}$ et $\underline{Q}$ où $\overline{Q}$ et $\underline{Q}$ sont les résultantes des forces agissant sur les points matériels respectivement des systèmes $\overline{\beta}$ et $\underline{\beta}$. Les systèmes $\overline{\alpha}$ et $\overline{\beta}$ étant équivalents, leurs résultantes $\overline{P}$ et $\overline{Q}$ sont confondues. De la même manière, les forces $\underline{P}$ et $\underline{Q}$ sont confondues. Ces forces étant confondues, leurs résultantes P et Q sont confondues elles aussi et, partant, leurs points d'application, i.e. les centres de gravité des systèmes α et β.

Remarquons que tout système de points matériels peut être remplacé par un système équivalent composé en tout et pour tout d'un point matériel. Pour cela il suffit de considérer le point matériel situé au centre de gravité du système initial et de masse égale à la masse totale de ce système. Ce point matériel constituera un système équivalent au système initial. Le point matériel situé au centre de gravité

d'un système et possédant une masse égale à la masse totale de ce système est appelé *centre matériel* du système considéré. On peut donc définir des systèmes équivalents comme des systèmes dont les centres matériels sont confondus.

§ 6 CENTRE DE GRAVITÉ D'UN SYSTÈME DE DEUX POINTS MATÉRIELS

Soit une barre de masse négligeable, portant deux charges P et Q distantes de d (fig. 25). On demande de trouver le centre de gravité du système ainsi constitué. En vertu de la proposition I du paragraphe précédent, ce problème est équivalent à celui qui consiste à déterminer sur quel point O doit reposer la barre pour être en équilibre. Désignons par u et v les distances respectives des charges P et Q au point cherché O. L'équilibre est réalisé si et seulement si le produit de la force P par le bras u est égal au produit de la force Q par le bras v, i.e.

$$Pu = Qv. \qquad (1)$$

En comparant la relation (1) à la relation

$$u + v = d, \qquad (2)$$

on tire

$$u = \frac{Q}{P+Q} d, \qquad v = \frac{P}{P+Q} d.$$

On remarque que $\frac{u}{v} = \frac{Q}{P}$, i.e. la longueur des bras est inversement proportionnelle aux charges.

En particulier, si les charges sont de même poids, le centre de gravité est situé au milieu de la barre (ce qui est évident et immédiat); inversement, si le centre de gravité se trouve au milieu de la barre, les deux charges sont égales.

Si l'une des charges est deux fois plus lourde que l'autre, le centre de gravité est plus proche de la charge la plus lourde et divise la barre dans le rapport 1:2; inversement, si le centre de gravité divise la barre dans le rapport 1:2,

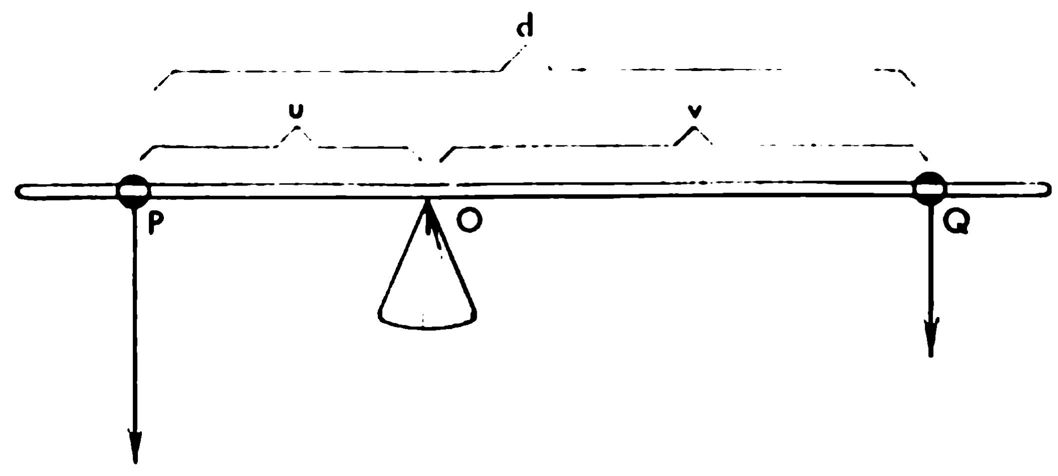


Fig. 25

l'une des charges (précisément celle qui est la plus proche du centre de gravité) est deux fois plus lourde que l'autre.

Comme le centre de gravité de la barre dont la masse est concentrée en M et N coïncide avec le centre de gravité du système des deux points matériels M et N (avec leurs masses respectives), nous déduisons le résultat suivant:

Le centre de gravité d'un système de deux points matériels est situé sur le segment reliant ces deux points. La distance de ce centre de gravité à ces points est inversement proportionnelle à leurs poids (donc à leurs masses).

En particulier, le centre de gravité est situé au milieu d'un segment si et seulement si ces deux points ont des masses égales; le centre de gravité divise un segment dans le rapport 1 : 2 si et seulement si l'un des points a une masse égale au double de celle de l'autre point (le centre est plus près du point le plus lourd).

§ 7 THÉORÈME RELATIF A L'INTERSECTION DES DROITES

Considérons un système quelconque α de points matériels, réunion des systèmes γ et δ. Désignons par M et N les centres matériels des systèmes γ et δ. En vertu de la proposition III du § 5, le système de points matériels M et N est équi-

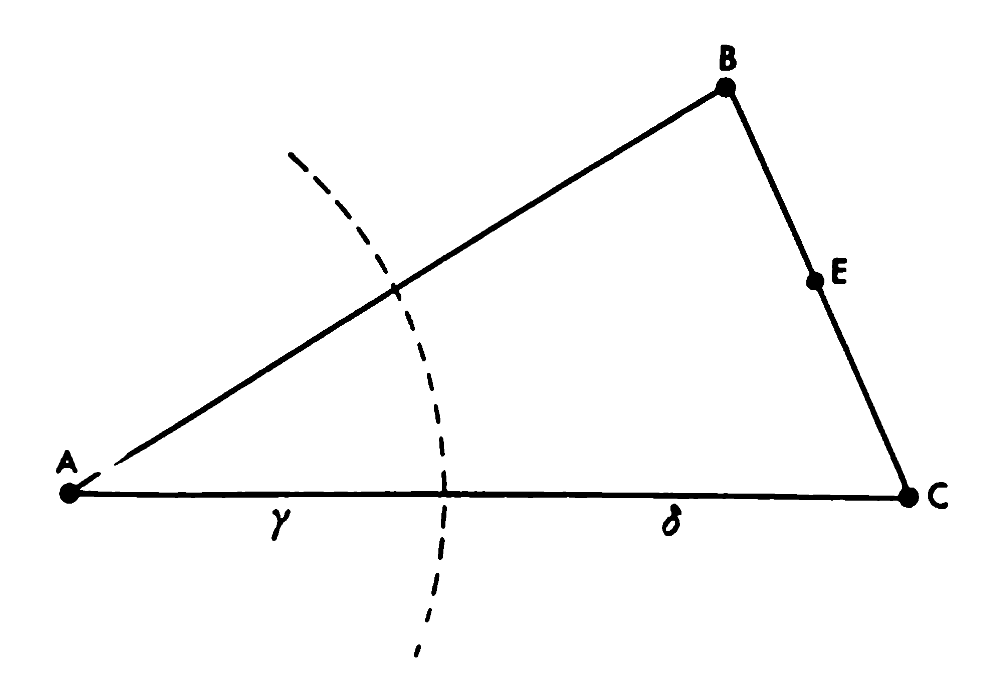


Fig. 26

valent au système α. Donc, le centre matériel du système α est confondu avec le centre matériel du système composé des points M et N et, de ce fait, est situé sur le segment reliant ces deux points. Comme le centre matériel du système se trouve au centre de gravité, nous déduisons le théorème suivant :

Le centre de gravité d'un système de points matériels, réunion des systèmes γ *et* δ, *est situé sur le segment reliant les centres de gravité des systèmes* γ *et* δ.

Nous allons voir maintenant donner trois applications de ce théorème en géométrie.

Les médianes d'un triangle sont concourantes. Considérons trois points matériels A, B et C de masse égale et disposons-les en triangle comme l'indique la fig. 26. Partageons ce système de points matériels en deux systèmes : l'un γ contenant le point A, l'autre δ contenant les points B et C. Le centre de gravité du système γ se trouve en A; celui du système δ au milieu E du côté BC en vertu des résultats du paragraphe précédent. En vertu du théorème que nous venons d'énoncer, le centre de gravité O du système de points A, B

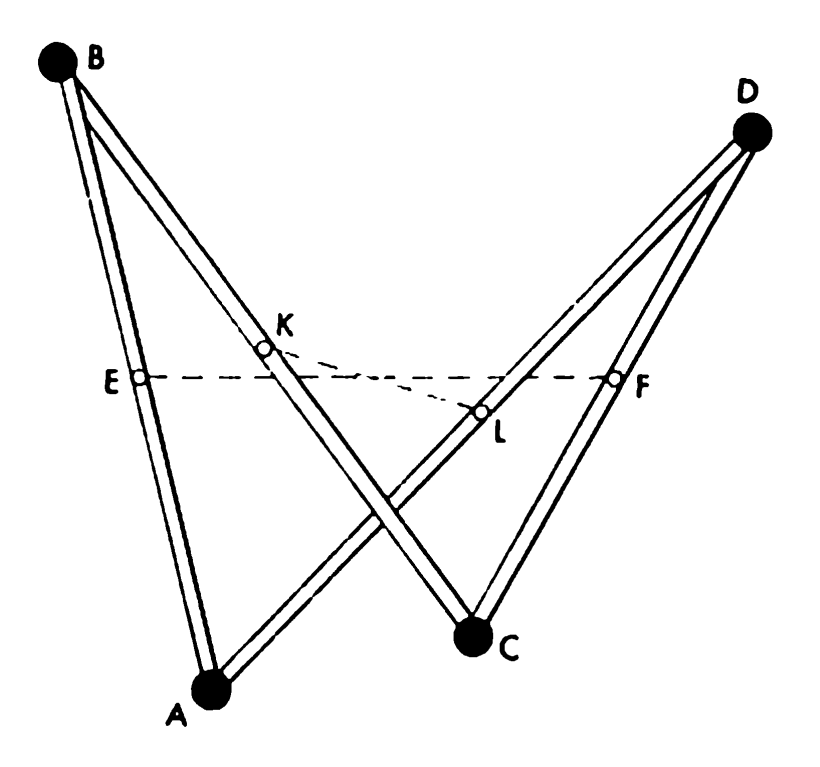


Fig. 27

et C se trouve sur la médiane AE. En reprenant ce même raisonnement, on établit que le point O est porté par les deux autres médianes. Donc les médianes du triangle ABC sont concourantes.

Comme le centre matériel M du système δ (ce centre est confondu avec E) est deux fois plus lourd que le centre matériel du système γ (ce centre est confondu avec le point A), le centre de gravité du système $\{M, A\}$ divise le segment EA dans le rapport 1:2 et, en outre, est plus proche de E. Or, ce centre de gravité n'est autre que le centre de gravité du système A, B et C, i.e. le point de concours des trois médianes. Nous avons déduit donc le théorème connu, à savoir que les médianes se coupent en un point situé au 2/3 de chacune d'elles à partir du sommet correspondant.

Quadrilatère gauche. Soit un quadrilatère gauche $ABCD$ (fig. 27). Montrons que les segments EF et KL qui joignent les milieux des côtés opposés se coupent. Disposons aux quatre sommets de ce quadrilatère des charges que nous assimilerons à des points matériels. Comme dans la « démonstra-

tion » du théorème des médianes, il nous suffit de prouver que le centre de gravité du quadrilatère est situé sur chacun des segments EF et KL. Commençons par le segment EF (pour KL la démonstration est analogue). Scindons le système de points A, B, C et D en deux systèmes : l'un contenant les points A et B, l'autre les points C et D. Le centre de gravité du premier système est situé en E, celui du second en F. En vertu du théorème formulé au début de ce paragraphe, le centre de gravité du quadrilatère se trouve sur le segment EF. Nous laissons au lecteur le soin de démontrer que les segments EF et KL se coupent en leurs milieux.

Théorème de Céva. Le théorème relatif à l'intersection des médianes d'un triangle est un cas particulier d'un théorème connu sous le nom de théorème de Céva :

Soient A_1, B_1 et C_1 *trois points situés respectivement sur les côtés* BC, CA *et* AB *d'un triangle* ABC (fig. 28); *la condition nécessaire et suffisante pour que les segments* AA_1, BB_1 *et* CC_1 *se coupent en un point est*

$$\frac{AB_1}{B_1C} \cdot \frac{CA_1}{A_1B} \cdot \frac{BC_1}{C_1A} = 1. \qquad (1)$$

Nous conseillons au lecteur de vérifier que la relation (1) est réellement satisfaite dans les cas suivants :

1. Les segments AA_1, BB_1 et CC_1 sont les bissectrices intérieures du triangle. (Indication : utiliser le théorème selon lequel la bissectrice intérieure d'un angle partage le côté opposé en segments proportionnels aux côtés adjacents.) D'où l'on déduira que les bissectrices du triangle sont concourantes.

2. Les segments AA_1, BB_1 et CC_1 sont les hauteurs du triangle. (Indication : utiliser la formule exprimant la longueur des segments déterminés sur le côté opposé par la hauteur.) On déduira ensuite que les hauteurs du triangle sont concourantes.

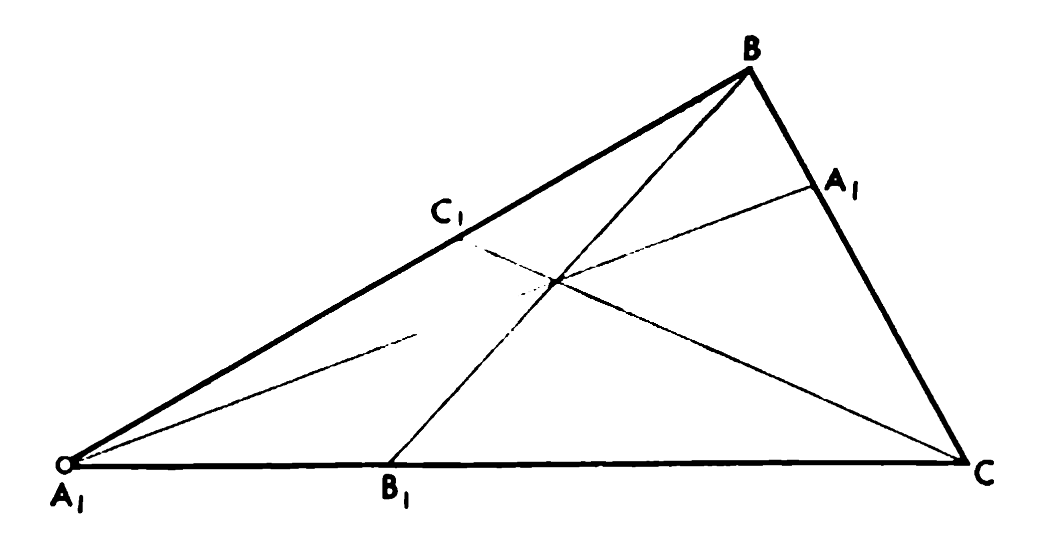


Fig. 28

Voyons maintenant la démonstration du théorème de Céva.

Supposons que la relation (1) est vérifiée et montrons que les segments AA_1, BB_1 et CC_1 concourent en un point. Le principe de la démonstration est le même que dans les deux exemples précédents. Nous allons montrer qu'avec des charges dûment choisies et disposées aux différents sommets, le centre de gravité du système se trouve sur chacun des segments AA_1, BB_1 et CC_1. Soient (à une échelle)

$$AB_1 = a, \qquad B_1C = b,$$
$$CA_1 = c, \qquad A_1B = d,$$
$$BC_1 = e, \qquad C_1A = f,$$

d'où

$$\frac{a}{b} \cdot \frac{c}{d} \cdot \frac{e}{f} = 1. \tag{2}$$

Plaçons en A, B et C des charges pesant respectivement bd, ca et ad (les unités sont quelconques). Le centre de gravité du système $\{A, C\}$ divise le côté AC en des segments inversement proportionnels aux charges disposées en A et C, i.e. il est confondu avec le point B_1. D'une façon analogue, le centre de gravité du système $\{B, C\}$ est confondu

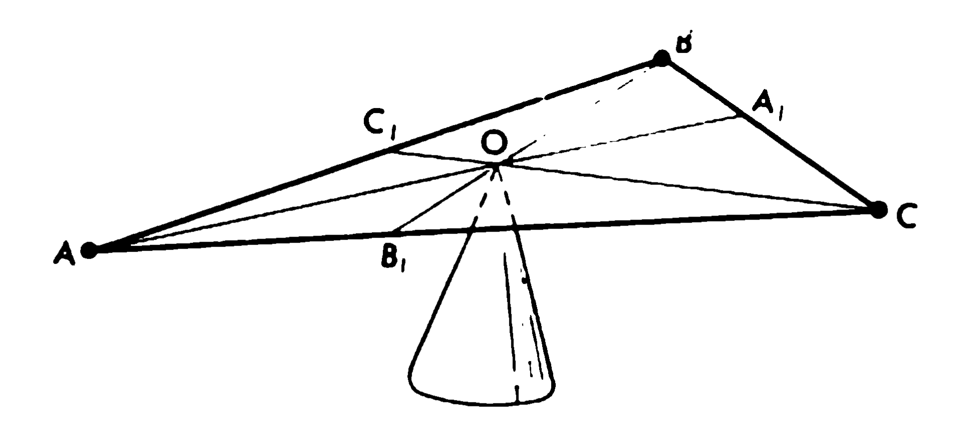


Fig. 29

avec le point A_1. Enfin, le centre de gravité du système $\{A, B\}$ est confondu avec le point dont les distances x et y aux sommets A et B sont inversement proportionnelles aux charges bd et ac, i. e.

$$\frac{x}{y} = \frac{ac}{bd} \quad \text{et} \quad x + y = f + e.$$

Or la relation (2) entraîne:

$$\frac{ac}{bd} = \frac{f}{e}.$$

Donc $x = f$ et $y = e$, et le centre de gravité du système $\{A, B\}$ est confondu avec le point C_1. Il reste à souligner qu'en vertu du théorème énoncé au début du paragraphe, le centre de gravité du triangle est situé sur chacun des segments AA_1, BB_1, CC_1.

Réciproquement, supposons que les segments AA_1, BB_1, CC_1 sont concourants en un point O, et montrons que la relation (1) est vérifiée. Pour cela assimilons notre triangle à une plaque de masse négligeable que nous posons sur une pointe (fig. 29). Désignons par O le point de contact. Disposons en A, B et C des charges propres à maintenir le triangle ABC en équilibre. En vertu de la proposition I du § 5, le point O sera le centre de gravité du triangle ABC avec ses charges, et, partant, le centre de gravité du système ainsi constitué. Le point O doit donc se trouver sur le segment qui joint A au centre de gravité du système $\{B, C\}$.

De la même façon on établit que B_1 et C_1 sont les centres de gravité des systèmes $\{C, A\}$ et $\{A, B\}$. Si on désigne les masses des points matériels A, B, C respectivement par p, q et r, en vertu du paragraphe précédent, on a:

$$\frac{AB_1}{B_1C}=\frac{r}{p}, \quad \frac{CA_1}{A_1B}=\frac{q}{r}, \quad \frac{BC_1}{C_1P}=\frac{p}{q}.$$

En multipliant membre à membre ces rapports on obtient la relation (1) cherchée.

§ 8 CENTRE DE GRAVITÉ D'UNE BARRE SUPPORTANT PLUSIEURS CHARGES

Généralisons les notions développées dans le § 6.

Imaginons-nous une barre de masse négligeable (fig. 30) supportant des charges P_1, P_2, . . ., P_n et proposons-nous d'en trouver le centre de gravité. Comme au § 6, il s'agit de trouver un point O tel que si la barre repose par ce point sur une pointe, elle soit en équilibre. Calculons l'abscisse x du point d'appui [1]).

Soit a_1 l'abscisse de la charge P_1, a_2 l'abscisse de la charge P_2, . . ., a_n l'abscisse de la charge P_n. Supposons par ailleurs (fig. 30) que le point d'appui O est compris entre P_k et P_{k+1} (il n'est pas exclu que le point O soit confondu avec le point d'application d'une charge). Les bras des forces P_1, . . ., P_k seront respectivement égaux à $x - a_1$, . . ., $x - a_k$ et les bras des forces P_{k+1}, . . ., P_n respectivement égaux à $a_{k+1} - x$, . . ., $a_n - x$. Comme la barre est en équilibre, la somme des moments des forces qui la font tourner dans le sens des aiguilles d'une montre

---

[1]) Ici et dans la suite on désignera par abscisse d'un point la distance de ce point à l'extrémité gauche de la barre.

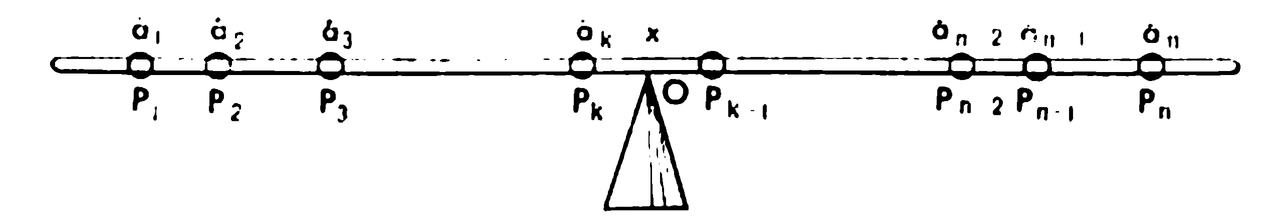


Fig. 30

est égale à la somme des moments des forces qui la font tourner dans le sens contraire, i. e.:

$$P_1(x - a_1) + \ldots + P_k(x - a_k) = P_{k+1}(a_{k+1} - x) + \\ + \ldots + P_n(a_n - x).$$

Groupons les termes en x dans le premier membre de l'équation et les autres dans le second membre:

$$(P_1 + \ldots + P_k + P_{k+1} + \ldots + P_n)\, x = \\ = P_1 a_1 + \ldots + P_k a_k + P_{k+1} a_{k+1} + \ldots + P_n a_n.$$

D'où

$$x = \frac{P_1 a_1 + \ldots + P_n a_n}{P_1 + \ldots + P_n}.$$

E x e m p l e 1. Supposons que $P_1 = 1$, $P_2 = 2$, ..., $P_n = n$; $a_1 = 1$, $a_2 = 2$, . . ., $a_n = n$. Il vient:

$$x = \frac{1^2 + 2^2 + \ldots + n^2}{1 + 2 + \ldots + n}.$$

Pour calculer la valeur de ce quotient, posons

$$1^3 = (1 + 0)^3 = 1^3,$$
$$2^3 = (1 + 1)^3 = 1^3 + 3 \cdot 1^2 + 3 \cdot 1 + 1,$$
$$3^3 = (2 + 1)^3 = 2^3 + 3 \cdot 2^2 + 3 \cdot 2 + 1,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$(n + 1)^3 = (n + 1)^3 = n^3 + 3 \cdot n^2 + 3 \cdot n + 1.$$

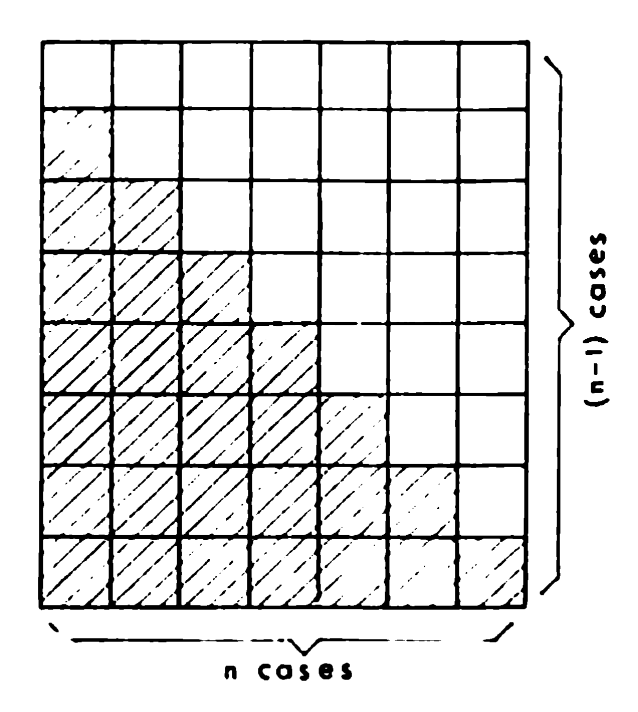


Fig. 31

En ajoutant membre à membre nous obtenons:

$$1^3 + 2^3 + \ldots + (n+1)^3 = 1 + (1^3 + 2^3 + \ldots + n^3) + \\ + 3(1^2 + 2^2 + \ldots + n^2) + 3(1 + 2 + \ldots + n) + n$$

ou

$$(n+1)^3 = 1 + 3(1^2 + 2^2 + \ldots + n^2) + \\ + 3(1 + 2 + \ldots + n) + n$$

ou encore

$$3(1^2 + 2^2 + \ldots + n^2) = (n+1)(n^2 + 2n) - \\ - 3(1 + 2 + \ldots + n).$$

Divisons les deux membres de cette égalité par $3(1 + 2 + \ldots + n)$:

$$\frac{1^2 + 2^2 + \ldots + n^2}{1 + 2 + \ldots + n} = \frac{(n+1)(n^2+2n)}{3(1+2+\ldots+n)} - 1.$$

La somme $1 + 2 + \ldots + n$ comporte autant d'unités que la fig. 31 comporte de cases hachurées. Or le nombre des cases hachurées est égal à celui des cases non hachurées

i. e. à la moitié des cases du rectangle, soit $\frac{n(n+1)}{2}$. Donc

$$x=\frac{1^2+2^2+\ldots+n^2}{1+2+\ldots+n}=\frac{(n+1)(n^2+2n)}{3(1+2+\ldots+n)}-1=$$
$$=\frac{(n+1)(n^2+2n)}{3\frac{n(n+1)}{2}}-1=\frac{2n+1}{3}.$$

D'où

$$1^2+2^2+\ldots+n^2=\frac{2n+1}{3}(1+2+\ldots+n)=$$
$$=\frac{2n+1}{3}\cdot\frac{n(n+1)}{2}=\frac{n(n+1)(2n+1)}{6}.$$

Exemple 2. Supposons que les charges de l'exemple 1, i. e. les charges de poids 1, 2, . . ., n, sont appliquées aux points d'abscisses 1^2, 2^2, . . ., n^2. Montrer alors que l'abscisse du centre de gravité est

$$\frac{n(n+1)}{2}.$$

(Indication. Se servir des résultats de l'exemple 1 et calculer la somme des cubes $1^3+2^3+\ldots+n^3$.)

Exemple 3. Supposons qu'aux points de l'exemple 1, i.e. aux points d'abscisses 1, 2, . . ., n, sont appliquées des charges de poids 1^2, 2^2, . . ., n^2. Montrer que l'abscisse du centre de gravité est

$$\frac{3n(n+1)}{2(2n+1)}.$$

§ 9 UN PROBLÈME DE LA THÉORIE DES NOMBRES (ÉNONCÉ)

Les notions faisant appel à la mécanique, outre des problèmes de géométrie, permettent quelquefois de résoudre des problèmes d'arithmétique. Et notamment celui, assez original, qui va suivre [1]).

[1]) Cet exercice nous a été aimablement communiqué par P. Novikov.

Soit une ligne de nombres positifs P_1, P_2, . . ., P_n; pour en réduire la longueur nous allons retrancher des nombres extrêmes P_1 et P_n le nombre P égal au plus petit d'entre eux, et ensuite ajouter P aux nombres moyens (s'il n'y a qu'un seul moyen, on lui ajoutera deux fois le nombre P). En outre (pourvu que $n > 2$), l'un des deux nombres extrêmes ou les deux à la fois (s'ils sont égaux) s'annulent, donc la nouvelle ligne sera plus courte que la ligne initiale. Si par exemple la ligne initiale est

$$1, 9, 3, 4,$$

on obtient

$$10, 4, 3.$$

Si la ligne initiale est

$$6, 2, 6,$$

la nouvelle ligne est composée d'un seul nombre, soit

$$14.$$

Toute nouvelle ligne sera par définition la « dérivée » de la précédente. Donc la ligne 10, 4, 3 est la dérivée de la ligne 1, 9, 3, 4; la ligne constituée du seul nombre 14, dérivée de la ligne 6, 2, 6. Remarquons que si une ligne est composée de deux termes, elle se confond avec sa dérivée.

Donnons-nous maintenant une ligne quelconque α. Calculons sa dérivée et désignons-la par α'. Prenons la dérivée de α' et désignons-la par α''. En poursuivant cette procédure, nous obtenons successivement les dérivées α', α'', . . ., α^n. Comme la longueur de la ligne est réduite à chaque dérivation d'un ou de deux termes, il arrivera un moment où nous obtiendrons une ligne qui renfermera 1 ou 2 termes; cette ligne sera précisément la dernière dérivée ou encore la caractéristique de la ligne. Ainsi la caractéristique de la ligne 1, 9, 3, 4 est la ligne 7, 10.

Observons que la somme des termes d'une ligne ne change lorsqu'on passe de cette ligne à sa dérivée. Donc, la

somme des termes de la caractéristique est égale à la somme des termes de la ligne initiale; en particulier, si la caractéristique ne renferme qu'un seul terme, celui-ci est précisément égal à la somme des termes de la ligne initiale. Lorsque la caractéristique est composée de deux termes, nous verrons dans le paragraphe suivant comment trouver les termes de la ligne initiale sans passer par le calcul des dérivées intermédiaires. Voyons maintenant les cas où la caractéristique est composée d'un seul terme et de deux termes. Essayons de tirer les choses au clair moyennant une « expérience de mathématiques ». Pour cela calculons les caractéristiques de quelques lignes de 1, 2, . . ., n termes. Nous écrirons les dérivées successives les unes sous les autres et soulignerons d'un trait le ou les termes moyens:

Exemple 1.
$$\begin{array}{ll} \alpha & 1\ \underline{2}\ 3 \\ \alpha' & \ \ 4\ 2 \end{array}$$

Exemple 2.
$$\begin{array}{ll} \alpha & 1\ \underline{2\ \ 3}\ 4 \\ \alpha' & \ \ 3\ \ \underline{4}\ 3 \\ \alpha'' & \ \ \ \ 10 \end{array}$$

Exemple 3.
$$\begin{array}{ll} \alpha & 1\ 2\ \underline{3}\ \ 4\ 5 \\ \alpha' & \ \ 2\ \underline{5\ \ 4}\ 4 \\ \alpha'' & \ \ \ \ 7\ \ \underline{6}\ 2 \\ \alpha''' & \ \ \ \ 5\ 10 \end{array}$$

Exemple 4.
$$\begin{array}{ll} \alpha & 1\ 2\ \underline{3\ \ 4}\ 5\ 6 \\ \alpha' & \ \ 2\ 4\ \ \underline{5}\ 5\ 5 \\ \alpha'' & \ \ \ \ 4\ \ \underline{9\ 5}\ 3 \\ \alpha''' & \ \ \ \ 1\ \underline{12}\ 8 \\ \alpha'''' & \ \ \ \ \ \ 14\ 7 \end{array}$$

Exemple 5.	α	1 2 3 4 5 6 7
	α'	2 3 6 5 6 6
	α''	3 8 7 6 4
	α'''	8 13 6 1
	α''''	7 14 7
	α'''''	28
Exemple 6.	α	1 2 3 4 5 6 7 8
	α'	2 3 5 6 6 7 7
	α''	3 5 10 6 7 5
	α'''	5 13 9 7 2
	α''''	3 13 13 7
	α'''''	16 16 4
	α''''''	12 24
Exemple 7.	α	1 2 3 4 5 6 7 8 9
	α'	2 3 4 7 6 7 8 8
	α''	3 4 9 8 7 8 6
	α'''	4 9 14 7 8 3
	α''''	1 9 17 10 8
	α'''''	9 19 10 7
	α''''''	2 26 17
	α'''''''	30 15

Les exemples traités montrent que la caractéristique d'un segment de la suite des nombres naturels peut renfermer un ou deux termes. Notons la règle suivante : la caractéristique est composée d'un seul terme si $n = 3k + 1$, dans tous les autres cas (i. e. lorsque $n = 3k$ ou $3k + 2$) elle renferme deux termes. Par ailleurs, si la caractéristique

est composée de deux termes l'un d'eux est le double de l'autre. Nous obtenons ainsi un curieux théorème:

La caractéristique de la ligne 1, 2, 3, . . ., *n est composée d'un terme si* $n = 3k + 1$ *et de deux termes si* $n = 3k$ ou $3k + 2$; *si elle est composée de deux termes l'un d'eux est le double de l'autre (plus exactement, le premier lorsque* $n = 3k$ *et le second lorsque* $n = 3k + 2$).

C'est ce théorème que nous allons démontrer dans le paragraphe suivant. Nous aurons en outre à l'esprit le problème suivant: pour quelles lignes la caractéristique est-elle composée d'un ou de deux termes et comment définir cela d'après les nombres $P_1, P_2, \ldots, P_n$ sans passer par le calcul des dérivées?

§ 10 UN PROBLÈME DE LA THÉORIE DES NOMBRES (RÉSOLUTION)

Abordons maintenant notre problème sous l'angle de la mécanique. Au lieu de la ligne des nombres $P_1, P_2, \ldots$ $\ldots, P_n$ considérons une barre (fig. 32,*a*) à laquelle sont appliquées en $A_1, A_2, \ldots, A_n$ (ces points sont tels que $A_1A_2 = A_2A_3 = \ldots = A_{n-1}A_n$) des charges $P_1, P_2, \ldots$ $\ldots, P_n$. A la dérivée de cette ligne correspond un nouveau système de charges, que nous appellerons système dérivé.

Montrons que le centre matériel du système de charges dérivé est confondu avec celui du système initial. Décomposons mentalement chacune des deux charges extrêmes P_1 et P_n en deux charges ponctuelles:

$$P_1 = P + (P_1 - P), \quad P_n = P + (P_n - P),$$

où P est la plus petite des charges P_1 et P_n (de sorte que l'une au moins des charges $P_1 - P$ et $P_n - P$ est nulle). Le système α peut alors être considéré comme la réunion de deux systèmes: un système $\underline{\alpha}$ composé des charges $P_1 - P$,

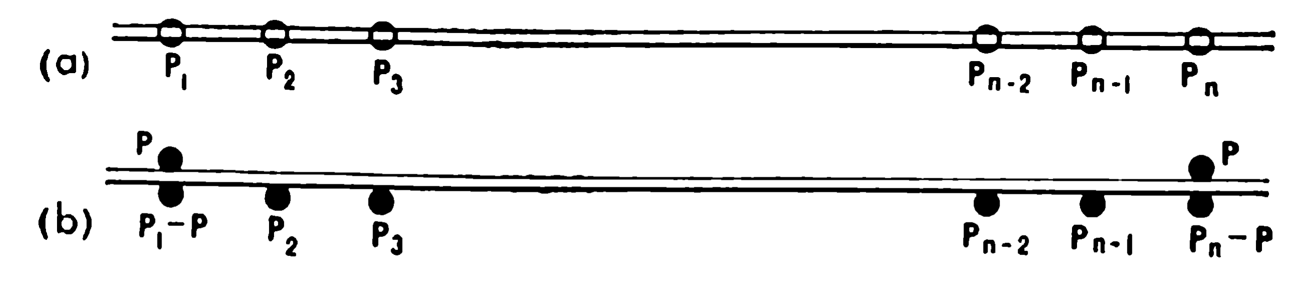


Fig. 32

$P_2, \ldots, P_{n-1}, P_n - P$ (ce système est représenté *sous* la barre sur la fig. 32,*b*) et un système $\overline{\alpha}$ composé de deux charges de poids P chacune (ce système est représenté *au-dessus* de la barre sur la fig. 32,*b*). Lorsqu'on passe au système dérivé, toutes les charges du système $\underline{\alpha}$ restent immobiles, quant aux charges du système $\overline{\alpha}$, elles se rapprochent à chaque dérivation d'une distance égale (pour s'ajouter aux termes moyens!). Autrement dit, le système dérivé α' est la réunion de deux systèmes $\underline{\alpha}'$ et $\overline{\alpha}'$ dont le premier est confondu avec le système $\underline{\alpha}$ et le second est déduit du système $\overline{\alpha}$ par un rapprochement de ses deux charges d'une distance égale. Les systèmes $\overline{\alpha}'$ et $\overline{\alpha}$ étant de toute évidence équivalents, il résulte, en vertu de la proposition III du § 5, que les systèmes α et α' sont aussi équivalents. Or, les centres matériels de deux systèmes équivalents sont confondus.

Ainsi donc, lorsqu'on passe d'un système de charges au système dérivé, le centre matériel (et par conséquent le centre de gravité) du système est invariable. Donc, il ne varie pas lorsqu'on passe du système initial à la caractéristique.

La caractéristique est donc composée d'une ou de deux charges ponctuelles admettant le même centre de gravité que le système initial des charges. Remarquons encore que les points d'application des charges ponctuelles qui constituent la caractéristique sont confondus avec les points

d'application de certaines charges du système initial, i. e. quelques-uns des points $A_1, \ldots, A_n$. Si la caractéristique est constituée de deux charges, celles-ci sont appliquées en deux points voisins A_j et A_{j+1}, et le centre de gravité est donc compris dans l'intervalle (A_j, A_{j+1}). La caractéristique est donc composée d'une seule charge si et seulement si son centre de gravité (ou, ce qui est équivalent, le centre de gravité du système initial) est confondu avec l'un quelconque des points $A_1, \ldots, A_n$.

En vertu du § 8, le centre de gravité du système initial admet pour abscisse

$$\frac{P_1a_1+P_2a_2+\ldots+P_na_n}{P_1+P_2+\ldots+P_n},$$

où $a_1, \ldots, a_n$ sont les abscisses des points $A_1, \ldots, A_n$. Si on fait $a_1 = 1, a_2 = 2, \ldots, a_n = n$, l'abscisse du centre de gravité est donnée par

$$\frac{P_1\cdot 1+P_2\cdot 2+\ldots+P_n\cdot n}{P_1+P_2+\ldots+P_n}.$$

La condition nécessaire et suffisante pour que le centre de gravité soit confondu avec l'un quelconque des points $A_1, A_2, \ldots, A_n$ est que son abscisse soit exprimée par un nombre entier. D'où nous déduisons: *la caractéristique de la ligne $P_1, P_2, \ldots, P_n$ est composée d'un seul terme si et seulement si le nombre* $\frac{P_1\cdot 1+P_2\cdot 2+\ldots+P_n\cdot n}{P_1+P_2+\ldots+P_n}$ *est entier.*

En particulier, si $P_1 = 1^2, P_2 = 2^2, \ldots, P_n = n^2$, alors (voir exemple 3 § 8) le nombre $\frac{P_1\cdot 1+P_2\cdot 2+\ldots+P_n\cdot n}{P_1+\ldots+P_n}$ est égal à $\frac{3n(n+1)}{2(2n+1)}$. Ce nombre n'est jamais entier (sauf dans le cas trivial où $n = 1$). En effet si ce nombre était entier, le nombre $4\frac{3n(n+1)}{2(2n+1)}$ serait entier, de même que la différence

$$4\frac{3n(n+1)}{2(2n+1)}-(3n+1)=\frac{n-1}{2n+1}.$$

Or, $\frac{n-1}{2n+1}$ n'est entier que si $n - 1 = 0$. Donc, lorsque $n > 1$, la caractéristique de la ligne 1, 4, 9, . . ., n^2 est toujours composée de deux termes.

Si $P_1 = 1$, $P_2 = 2$, . . ., $P_n = n$, alors (voir exemple 1 § 8) $\frac{P_1 \cdot 1 + \ldots + P_n \cdot n}{P_1 + \ldots + P_n} = \frac{2n+1}{3}$. Si $n = 3k + 1$, alors $\frac{2n+1}{3} = 2k + 1$ est entier et la caractéristique est composée d'un seul terme. Si $n = 3k$, alors $\frac{2n+1}{3} = 2k + \frac{1}{3}$; si $n = 3k + 2$ alors $\frac{2n+1}{3} = (2k + 1) + \frac{2}{3}$; dans les deux derniers cas, le nombre $\frac{2n+1}{3}$ n'est pas entier et la caractéristique est donc composée de deux termes. Ceci achève la démonstration de la première partie du théorème énoncé à la fin du paragraphe précédent.

Si $n = 3k$, le centre de gravité du système de charges 1, 2, . . ., n admet, comme nous venons de le voir, pour abscisse le nombre $2k + \frac{1}{3}$. Il est donc compris entre les points A_{2k} et A_{2k+1}. Ces points sont donc précisément les points d'application des charges qui constituent la caractéristique du système initial. Comme le centre de gravité s'est avéré être deux fois plus près de A_{2k} que de A_{2k+1}, on en déduit que la charge appliquée en A_{2k} est deux fois plus lourde que celle appliquée en A_{2k+1}. Donc, lorsque $n = 3k$, le premier terme de la caractéristique est le double du second. Si maintenant $n = 3k + 2$ l'abscisse du centre de gravité est égale à $(2k + 1) + 2/3$. Dans ce cas, le centre de gravité est compris entre les points A_{2k+1} et A_{2k+2}, deux fois plus près du second point. Donc la charge (de la caractéristique) qui est appliquée au deuxième point est deux fois plus lourde que celle appliquée en A_{2k+1}; autrement dit, le second terme de la caractéristique est le double du premier. Ce qui achève la démonstration du théorème du § 9.

La somme des termes de la caractéristique de la ligne 1, 2, . . ., n est égale à la somme des termes de cette ligne, soit $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$ (voir exemple 1 § 8). Sachant que l'un des termes est le double de l'autre, on peut les calculer tous deux. Nous obtenons les valeurs suivantes: $\frac{n(n+1)}{6}$ pour le plus petit et $\frac{n(n+1)}{3}$ pour le plus grand.

Finalement, la caractéristique de la ligne 1, 2, . . ., n s'écrit

$$\frac{n(n+1)}{3}, \qquad \frac{n(n+1)}{6} \qquad (\text{si } n = 3k),$$

$$\frac{n(n+1)}{2} \qquad\qquad (\text{si } n = 3k+1),$$

$$\frac{n(n+1)}{6}, \qquad \frac{n(n+1)}{3} \qquad (\text{si } n = 3k+2).$$

Dans le cas général, pour calculer la caractéristique de la ligne $P_1, \ldots, P_n$ il importe de procéder comme suit. Il y a lieu de calculer tout d'abord le nombre

$$x = \frac{P_1 \cdot 1 + P_2 \cdot 2 + \ldots + P_n \cdot n}{P_1 + P_2 + \ldots + P_n}.$$

Si x est entier, la caractéristique est composée d'un seul terme égal à la somme $P_1 + P_2 + \ldots + P_n$. Si x n'est pas entier, il faut alors le mettre sous la forme:

$$x = y + z,$$

où y est un entier et $0 < z < 1$. Dans ce cas, la caractéristique est composée de deux termes Q_1 et Q_2 tels que

$$Q_1 + Q_2 = P_1 + P_2 + \ldots + P_n. \tag{1}$$

En vertu du § 5, entre les nombres Q_1 et Q_2 on a la relation

$$\frac{Q_1}{Q_2} = \frac{1-z}{z}. \tag{2}$$

puisque Q_1 et Q_2 peuvent être considérés comme des charges appliquées aux points d'abscisses y et $y + 1$ et dont le centre de gravité

est d'abscisse x. En résolvant simultanément les équations (1) et (2), on trouve

$$Q_1 = \frac{1-z}{z} Q_2, \quad \frac{1}{z} Q_2 = \frac{1-z}{z} Q_2 + Q_2 = Q_1 + Q_2 = P_1 + \ldots + P_n,$$

d'où

$$Q_2 = (P_1 + \ldots + P_n) z, \quad Q_1 = (P_1 + \ldots + P_n)(1-z). \tag{3}$$

Clôturons ce paragraphe par la recherche de la caractéristique de la ligne 1, 4, 9, ..., n^2 ($n > 1$). Nous savons que dans ce cas $x = \frac{3n(n+1)}{2(2n+1)}$ et la caractéristique comporte deux termes. Pour les calculer il nous faut d'abord connaître z. Posons à ces fins $n = 4k + r$ où $r = 0, 1, 2, 3$. Il vient alors :

$$x = \frac{3n(n+1)}{2(2n+1)} = \frac{3(4k+r)(4k+r+1)}{2[2(4k+r)+1]} = 3k + \frac{3}{2}\,\frac{2k+4kr+r+r^2}{8k+2r+1}.$$

Si $r = 0$, alors $x = 3k + \frac{3}{2}\,\frac{2k}{8k+1} = 3k + \frac{3k}{8k+1}$; donc $z = \frac{3k}{8k+1} = \frac{v}{2n+1}$, où $v = 3k$.

Si $r = 1$, alors $x = 3k + \frac{3}{2}\,\frac{6k+2}{8k+3} = 3k + 1 + \frac{k}{8k+3}$, donc $z = \frac{k}{8k+3} = \frac{v}{2n+1}$, où $v = k$.

Si $r = 2$, alors $x = 3k + \frac{3}{2}\,\frac{10k+6}{8k+5} = 3k + 1 + \frac{7k+4}{8k+5}$, donc $z = \frac{7k+4}{8k+5} = \frac{v}{2n+1}$, où $v = 7k + 4$.

Si $r = 3$, alors $x = 3k + \frac{3}{2}\,\frac{14k+12}{8k+7} = 3k + 2 + \frac{5k+4}{8k+7}$, donc $z = \frac{5k+4}{8k+7} = \frac{v}{2n+1}$, où $v = 5k + 4$.

La somme des termes de la caractéristique est égale à $1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$ (voir exemple 1 du § 8). Les termes de

la caractéristique sont donnés par les formules (3):

$$Q_2 = \frac{n(n+1)(2n+1)}{6} z = \frac{n(n+1)(2n+1)}{6} \frac{v}{2n+1} = \frac{n(n+1)v}{6},$$

$$Q_1 = \frac{n(n+1)(2n+1)}{6}(1-z) = \frac{n(n+1)(2n+1)}{6} \frac{2n+1-v}{2n+1} =$$

$$= \frac{n(n+1)(2n+1-v)}{6}.$$

En remplaçant n par $4k$, $4k+1$, $4k+2$, $4k+3$ et compte tenu de la dépendance de v et de k, on trouve en définitive que la caractéristique de la ligne $1, 4, 9, \ldots, n^2$ est de la forme:

$$\frac{2k(4k+1)(5k+1)}{3}, \quad 2k^2(4k+1) \quad (\text{si } n = 4k),$$

$$\frac{(2k+1)(4k+1)(7k+3)}{3}, \quad \frac{k(2k+1)(4k+1)}{3} \quad (\text{si } n = 4k+1),$$

$$\frac{(k+1)(2k+1)(4k+3)}{3}, \quad \frac{(2k+1)(4k+3)(7k+4)}{3} \quad (\text{si } n = 4k+2),$$

$$2(k+1)^2(4k+3), \quad \frac{2(k+1)(4k+3)(5k+4)}{3} \quad (\text{si } n = 4k+3).$$

§ 11 L'IMPOSSIBILITÉ DE CONSTRUIRE UN MOTEUR PERPÉTUEL

Considérons un polygone convexe quelconque et un point O contenu à l'intérieur. Abaissons de O les perpendiculaires sur les côtés du polygone. Les pieds de ces perpendiculaires peuvent être situés aussi bien sur les côtés (telle la perpendiculaire abaissée sur AB, voir fig. 33) que sur leurs prolongements (telles les perpendiculaires abaissées sur FA et EF). Il peut bien sûr arriver que les pieds de toutes les perpendiculaires tombent sur les côtés (il en sera ainsi si le point O est le centre d'un polygone régulier). Une question se pose: est-il possible que les pieds de toutes les perpendiculaires tombent non pas sur les côtés mais sur leurs prolongements? On démontre qu'il ne peut en être ainsi, i.e. *le pied d'au moins une perpendiculaire est*

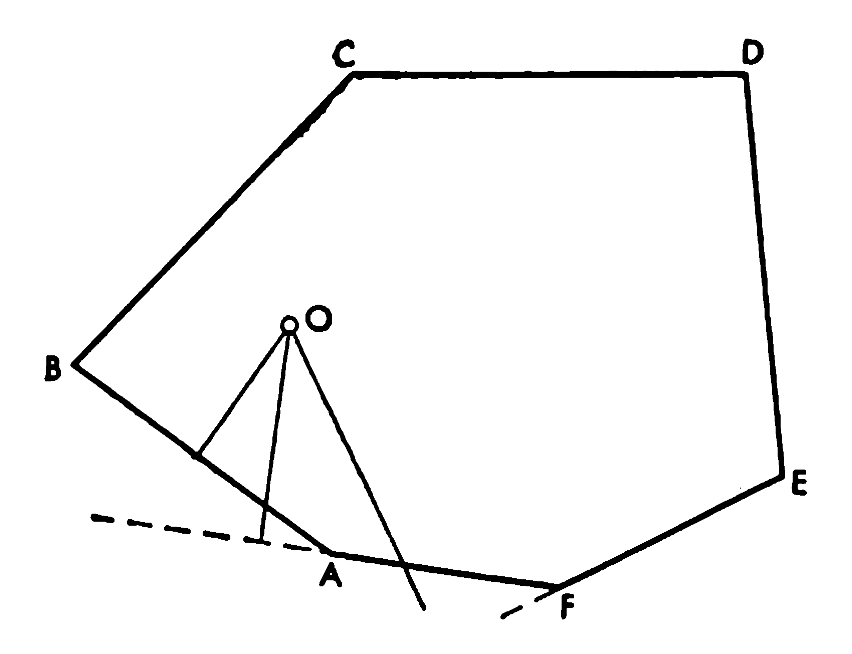


Fig. 33

situé sur le côté correspondant (et non sur son prolongement). En vue de démontrer cette assertion nous allons nous servir du principe formulé dans le titre de ce paragraphe. Assimilons le polygone à une lamelle dont le centre de gravité est confondu avec le point O (on peut, par exemple, considérer que la lamelle est de masse négligeable et qu'en O est appliquée une charge ponctuelle). Faisons reposer la lamelle par l'une de ses arêtes sur un plan horizontal. Si nous la posons verticalement, elle ne tombera pas; on peut à la rigueur la maintenir des deux côtés par deux tiges horizontales (voir fig. 34). Alors, soit le polygone reste au repos, soit il se met en mouvement: s'il se met en mouvement, il finira tôt ou tard par s'immobiliser (sinon le mouvement serait perpétuel). Dans cette position la perpendiculaire abaissée du centre de gravité sur l'arête par laquelle le polygone repose sur le plan, coupe obligatoirement cette arête (en vertu de la proposition II du § 5). Donc nous avons défini une arête contenant le pied de la perpendiculaire abaissée de O sur elle.

Un théorème analogue est valable pour le polyèdre. Plus exactement, soit un polyèdre convexe et un point O

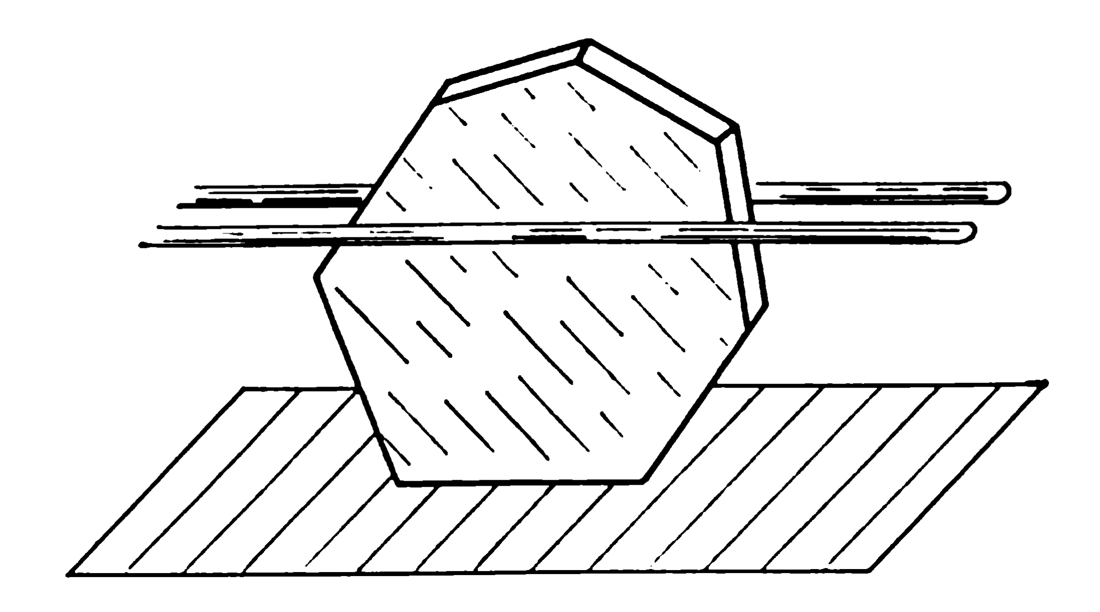


Fig. 34

contenu à l'intérieur de ce dernier. A partir de O abaissons des perpendiculaires sur toutes les faces du polyèdre. Alors *le pied d'au moins une perpendiculaire est situé sur une face (et non sur son prolongement).*

Pour démontrer ce théorème assimilons ce polyèdre à un solide admettant le point O pour centre de gravité. (Nous considérerons par exemple que le polyèdre est impondérable et qu'en O est appliquée une charge ponctuelle.) Faisons maintenant reposer notre polyèdre par l'une quelconque de ses faces sur un plan horizontal. Alors, soit il restera au repos, soit il se mettra en mouvement et tôt ou tard finira par s'immobiliser. Il existe donc une position dans laquelle le polyèdre est en équilibre sur le plan. Considérons la face par laquelle il repose sur ce plan. En vertu de la proposition II du § 5, la perpendiculaire abaissée du point O sur cette face coupe cette dernière. On a donc trouvé une face satisfaisant aux conditions du théorème formulé.

CONCLUSION

Tous les problèmes que nous avons traités auraient pu être résolus par des méthodes purement mathématiques.

Il faut toutefois se garder de penser que l'introduction de notions de mécanique dans la résolution de problèmes de mathématiques est un simple jeu d'esprit. Ces méthodes ont une signification historique et pratique.

Archimède déjà utilisait les lois de l'équilibre pour déterminer l'aire de figures curvilignes; ainsi pour calculer l'aire d'un segment de parabole (on appelle segment de parabole la figure limitée par un arc de parabole et la corde qui l'intercepte) il « suspendit » ce segment à un bras de levier. Tout au long des siècles la mécanique et la physique ont exercé et continuent d'exercer une influence indiscutable sur la résolution de nombreux problèmes (à défaut de fournir une démonstration rigoureuse elles mettent sur la voie).

En mathématiques, l'usage des lois de la mécanique représente un cas particulier d'une méthode générale consistant à utiliser dans le sens inverse les liens existant entre les phénomènes de la nature et leur description mathématique. Décrire un phénomène en mathématiques, c'est établir les formules permettant de calculer les caractéristiques physiques de ce phénomène (vitesse, température, distance, etc.) ou encore les équations dont les solutions sont ces caractéristiques. Utiliser directement la description mathématique c'est déterminer les valeurs numériques des caractéristiques sans observer ce phénomène, mais en procédant aux calculs indispensables grâce aux formules ou aux équations établies. Cette marche est à conseiller notamment lorsque le phénomène est complexe et sa description simple. On peut néanmoins suivre la marche inverse: au lieu de calculer la valeur des formules ou résoudre les équations on peut réaliser ce phénomène en laboratoire, mesurer les caractéristiques et, ce faisant, établir empiriquement les valeurs des formules et des solutions des équations. Cette méthode porte ses fruits lorsque nous nous trouvons en présence de formules compliquées ou d'équations difficilement résolubles décrivant un phénomène dont la réalisation en laboratoire est relativement aisée.

Ce principe est posé à la base de la simulation (i. e. l'étude des phénomènes physiques sur des modèles ou simulateurs) qui se réalise suivant le schéma suivant : soit à étudier un phénomène A (c'est le phénomène simulé); celui-ci est décrit par des formules et des équations E qui servent simultanément à la description d'un phénomène A' (c'est le phénomène simulatif) réalisable en laboratoire; en réalisant A', on détermine les valeurs des formules ou des solutions de l'équation E et par suite on établit les propriétés nécessaires du phénomène A. On dira alors à propos du phénomène A (de même d'ailleurs que des formules et équations E) qu'il *est simulé* par le phénomène A'. Ajoutons encore que si deux phénomènes sont décrits par les mêmes formules ou équations, on dit alors qu'il existe une *analogie* entre ces phénomènes.

Un exemple simple de simulation nous est fourni par la résolution au § 4 du problème de Steinhauss. Le phénomène simulé A est la marche des enfants vers l'école; la caractéristique qui nous intéresse est le temps mis pour se rendre à l'école. Ce temps représente la valeur d'une certaine formule. On trouve le phénomène simulatif A' (suspension des charges) pour lequel la valeur de cette formule correspond à une valeur de l'énergie potentielle. Le phénomène A' est réalisable expérimentalement. La simulation est d'une grande importance dans la pratique (les phénomènes simulatifs sont essentiellement des phénomènes hydrodynamiques et électriques). En particulier, la simulation est à la base de nombreux intégrateurs dont le plus simple, l'intégrateur à friction, va être brièvement examiné.

Soient une droite p et une courbe l ne possédant pas de point commun avec cette droite (fig. 35) et telle que toute perpendiculaire à p coupe l en un point au plus. Soient AB et CD deux perpendiculaires à p. On se propose de calculer l'aire du trapèze curviligne $ABCD$. A ces fins partageons l'intervalle AC en n intervalles partiels à l'aide des points

$$X_1 = A, X_2, X_3, \ldots, X_n, X_{n+1} = C$$

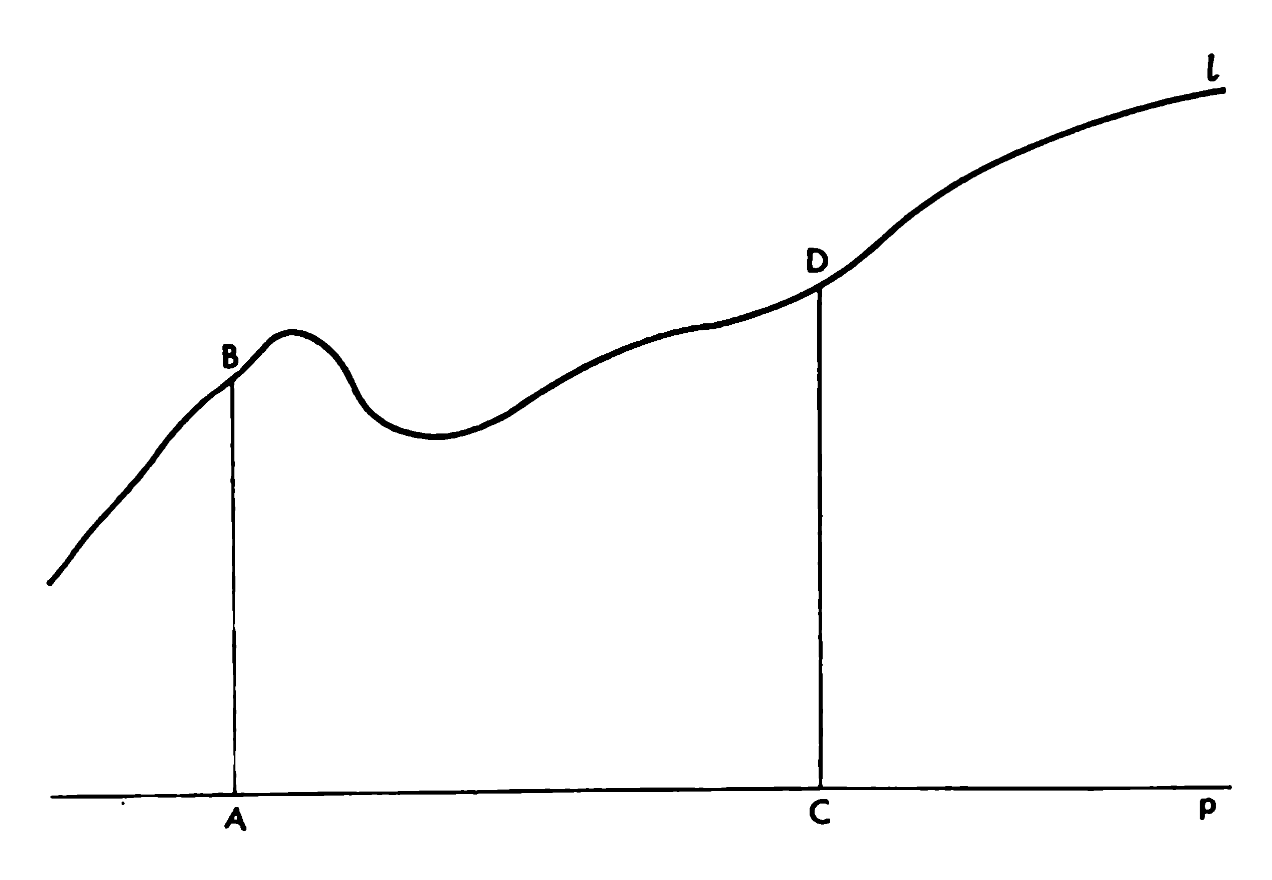


Fig. 35

(sur la fig. 36 on a $n = 6$). Dans chaque intervalle partiel $X_i X_{i+1}$ prenons un point T_i à partir duquel nous élevons une perpendiculaire qui coupera la courbe l en Y_i. A partir de chaque point Y_i menons une parallèle à p jusqu'à son intersection avec les perpendiculaires élevées des points X_i. L'aire S' de la figure hachurée (fig. 36) est sensiblement égale à l'aire S du trapèze curviligne, et en outre S' et S sont d'autant plus voisines que les intervalles partiels $X_i X_{i+1}$ sont plus nombreux. Supposons que les intervalles partiels $X_1 X_2, \ldots, X_n X_{n+1}$ sont respectivement de longueur $\alpha_1, \ldots, \alpha_n$ et les segments $T_1 Y_1, \ldots, T_n Y_n$ respectivement de $y_1, \ldots, y_n$. L'aire S' est

$$\alpha_1 y_1 + \ldots + \alpha_n y_n. \tag{1}$$

Considérons maintenant un système composé du disque

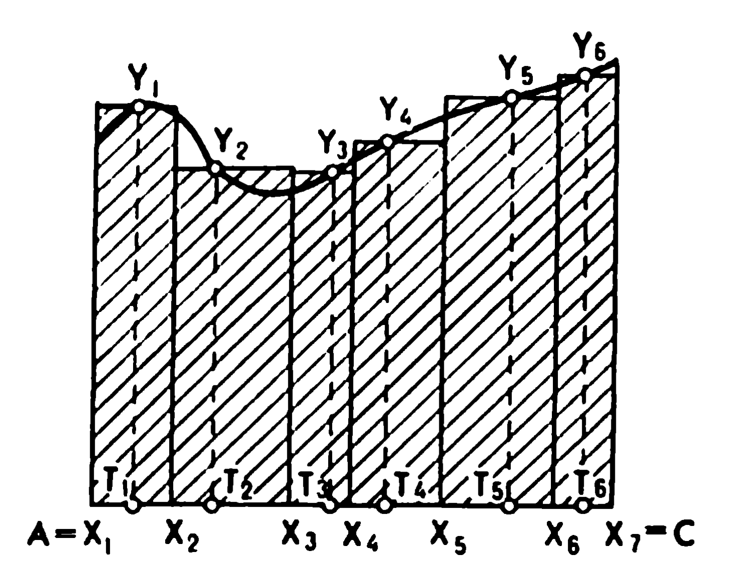


Fig. 36

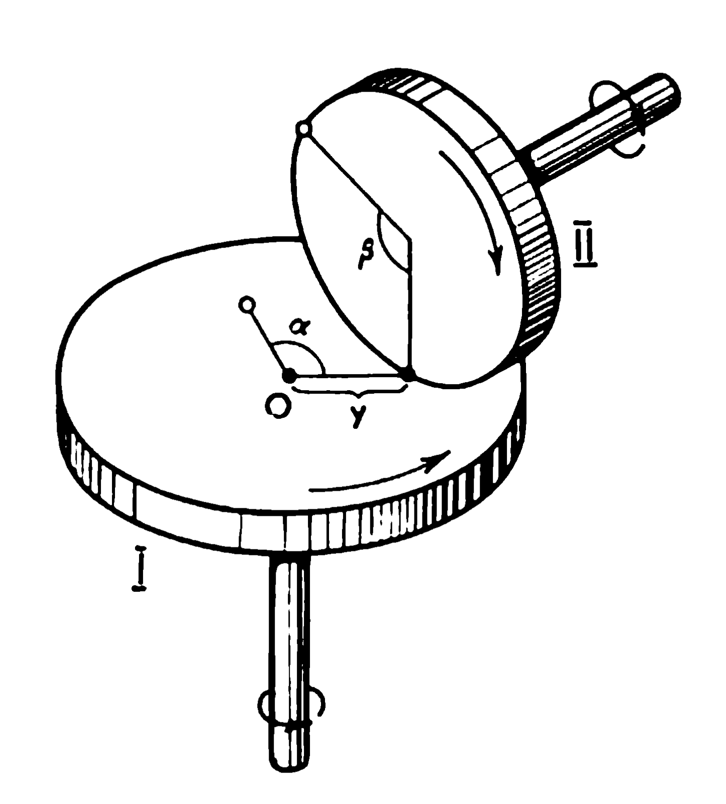


Fig. 37

I et du disque II reposant sur le premier (fig. 37). Le disque II peut se déplacer le long de son axe. Supposons qu'à un moment donné la distance du disque II au centre O du disque I est y. Faisons subir au disque I une rotation d'angle α (en radians). Le point par lequel le disque I était en con-

tact avec le disque II décrit un arc égal à αy ; c'est la distance que parcourra également le point par lequel le disque II reposait sur le disque I. Si le rayon du disque II est r et β l'angle dont il a pivoté, on aura donc la relation $\alpha y = \beta r$. Posons $r = 1$, alors $\beta = \alpha y$. Si maintenant on place le disque II à une distance y_1 du centre O du disque I et qu'on fasse pivoter ce dernier d'un angle α_2 et qu'ensuite on recommence n fois cette opération en faisant varier à chaque fois l'angle α_n et la distance y_n, alors l'angle total dont aura tourné le disque II représentera précisément la valeur numérique de la quantité (1), i.e. l'aire S' de la figure hachurée (voir fig. 36).

Traçons une droite p et une courbe l sur une feuille de papier (fig. 38). Considérons le dispositif suivant. Une règle est fixée par ses extrémités à deux roues dont l'une est astreinte à se déplacer le long de la droite p (de sorte que la règle sera toujours perpendiculaire à p). La règle est munie d'un curseur pouvant glisser le long d'elle, et portant en son centre une pointe en contact avec la feuillle de papier. Supposons que la règle est liée avec le disque I et le disque II de la manière suivante: la rotation de la roue à laquelle est fixée la règle entraîne une rotation d'un même angle du disque I ; la distance du disque II au centre O du disque I est égale à tout instant à la distance du centre du curseur à l'extrémité de la règle qui se déplace le long de p. Prenons le rayon des roues de la règle égal à 1. Faisons maintenant suivre à la pointe du curseur le contour de la figure hachurée représentée sur la fig. 38. Ceci étant, la règle se déplacera le long de la droite p et le curseur glissera le long de la règle. De ce que nous avons dit dans l'alinéa précédent, il résulte que le disque II tournera d'une quantité égale à la mesure de l'aire S'. Plus l'intervalle AC est subdivisé en intervalles partiels moindres, plus la ligne polygonale qui limite la figure hachurée se rapprochera de la courbe l et plus l'aire S' sera voisine de S. Si donc la pointe du curseur suit le tracé de la courbe l, le disque II

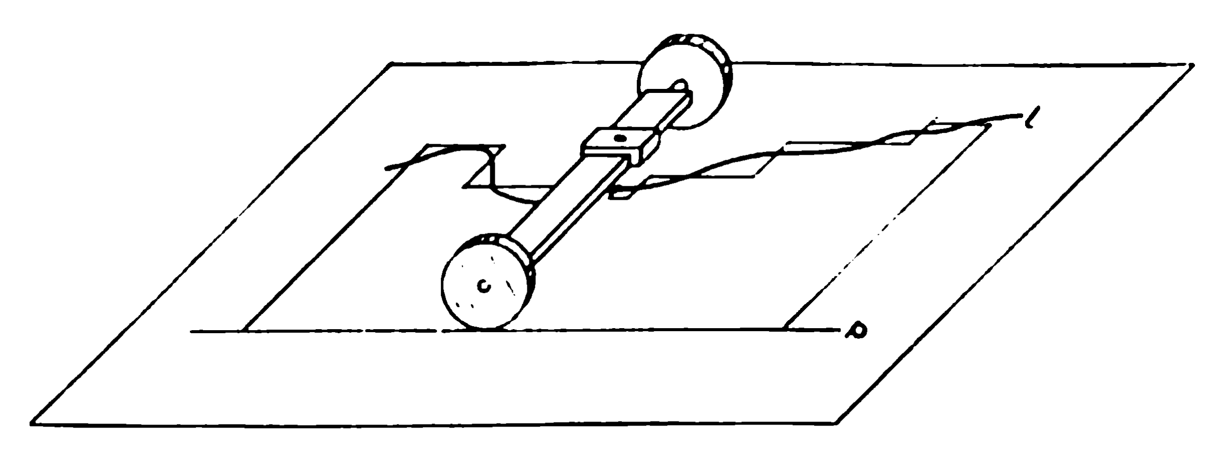


Fig. 38

tournera d'une quantité égale à la mesure de l'aire S. L'angle de rotation du disque II peut facilement être enregistré.

Le processus simulé ici est la mesure de l'aire du trapèze curviligne par le calcul des aires des figures hachurées. Ce processus est décrit par la formule (1). Cette même formule, on l'a vu, décrit un autre processus lié au mécanisme à friction représenté sur la fig. 37. Ceci nous a permis, au moyen du mécanisme indiqué, de simuler la formule (1) et la mesure de l'aire d'une surface.

Les analogies mécaniques nous ont conduits tout naturellement aux appareils destinés aux calculs des aires de figures curvilignes (en vérité, de figures de forme particulière, plus exactement de trapèzes curvilignes).

Pour conclure cet exposé faisons une remarque de principe. Les raisonnements fondés sur des notions de mécanique peuvent paraître incorrects à tout esprit chicanier. Les objections peuvent être formulées dès l'exposé du premier exemple au § 1. En effet, un fil est loin d'être une droite (d'abord il est d'une certaine grosseur et ensuite un examen au microscope en fait ressortir toutes les difformités), de sorte que lorsqu'on dit « le fil se tend suivant une droite » on est loin du sens exact. D'autre part la rigueur mathématique interdit de parler de l'extrémité d'un fil comme d'un point (de même d'ailleurs que d'affirmer qu'une charge est attachée à un fil en un point défini). Enfin, tout fil réel est extensible en sorte que sa longueur au repos ou avec une charge n'est pas la même.

Les critiques de cette nature seraient nombreuses. Néanmoins, il faut se garder de penser que les raisonnements développés sont

sans valeur. Pour notre part, nous les mettons sur un même rang que les démonstrations géométriques pures. En effet, nos raisonnements sur les charges, les fils, etc., revêtent tous un caractère approximatif et ils sont d'autant plus rigoureux que sont idéales les propriétés des objets étudiés (fil, charges, etc.). En ce qui nous concerne, certaines propositions (par exemple « le fil se tend suivant une droite ») sont d'autant plus vraies que le fil est fin ; on dit alors que cette proposition est vraie pour un fil infiniment fin. Les autres propositions (par exemple « la longueur du fil est la même avec ou sans charge ») sont d'autant plus vraies que le fil est inextensible ; on dit alors que cette proposition est vraie pour un fil absolument inextensible. Il va de soi qu'un fil infiniment fin ou absolument inextensible n'existe pas dans la nature. C'est une notion idéale tout comme est idéale la notion de point d'application d'une force étudiée en physique élémentaire ; les conclusions tirées de l'étude des forces et de leurs points d'applications sont d'autant plus réelles qu'est moindre l'aire de l'interaction des solides. On a également idéalisé la notion de point matériel ; tout ce qui a été établi pour les points matériels et généralisés aux solides est d'autant plus vrai que les volumes de ces solides sont petits.

Qu'il nous soit permis de souligner que toutes les notions et lois de la mécanique (et d'une façon générale de la physique) ont été établies à partir de phénomènes de la nature idéalisés ; aussi le lycéen est-il constamment confronté à des idéalisations de ce genre, qu'il le veuille ou non. Prenons par exemple une des premières notions de la mécanique : le mouvement uniforme. Il n'existe pas dans la nature de mouvements uniformes à proprement parler (mêmes les aiguilles d'une montre sont animées, à un examen minutieux, d'un mouvement non uniforme). Cependant, cette non-uniformité est si insignifiante qu'on peut la négliger ; pourtant, la notion de mouvement uniforme existe ! On peut en dire autant de la première loi de Newton : en effet il n'existe de corps qui ne soit pas soumis à l'action d'un autre corps. Les exemples de ce genre abondent en physique.

Nous remarquons donc que la mécanique étudie en fait non pas des solides et des processus réels, mais des solides et des processus idéalisés tels « un solide soustrait à l'action d'autres solides », « un mouvement rectiligne uniforme » etc. Et répétons-le, nous agissons avec ces objets idéalisés comme s'ils étaient réels, i.e. comme s'ils existaient dans la nature.

Soulignons enfin qu'en mécanique les notions abstraites (point matériel, fil impondérable infiniment fin et absolument inextensible, etc.) ne se distinguent en rien des notions abstraites dont use la géométrie (exemple le point, la droite, le plan, etc.) ; en effet, point, droite, plan, etc., n'existent pas dans la nature sous la forme d'objets réels et les propositions faites à leur sujet sont d'autant plus vraies que leurs propriétés approchent celles du point, du plan, etc.

Y. Lioubitch, L. Chor

Méthode cinématique dans les problèmes de géométrie

En feuilletant un sérieux livre de mathématiques, il nous a été donné de tomber sur un problème tout Conan Doyle ou Stevenson. Il s'agissait de trouver un trésor enfoui dans un endroit où se dressaient trois arbres : un chêne, un pin et un bouleau. Les indications étaient les suivantes : se placer sous le bouleau (désigné par la lettre B sur la fig. 1) face au chêne et au pin (représentés respectivement par les lettres C et P sur la fig. 1). Le chêne sera situé alors à droite et le pin à gauche. Se rendre jusqu'au chêne en comptant le nombre de pas ; une fois le chêne atteint, tourner de 90° à droite, faire autant de pas que du bouleau au chêne et marquer cet endroit d'un jalon J_1. Retourner ensuite au bouleau puis au pin en comptant toujours le nombre de pas. Une fois arrivé au pin, tourner de 90° à gauche et compter autant de pas que du bouleau au pin, et poser le deuxième jalon J_2. Le trésor est enfoui entre les deux jalons, exactement au milieu (le point T sur la fig. 1).

Avec des indications aussi précises, direz-vous, il n'est pas besoin d'être sorcier pour trouver la cachette du trésor. Et pourtant... Lorsque le chercheur se présenta sur les lieux, il découvrit bien le chêne, le pin, mais point de bouleau. Il réussit néanmoins à déterrer le trésor. On demande comment s'y est-il pris?

Comme ce problème figurait dans un Sérieux Livre de Mathématiques, il faut croire que sa solution ne procédait pas du hasard. Et en effet, ce problème admet une solution qui de surcroît est accessible au lycéen.

A partir des points J_1, J_2, B et T abaissons sur la droite CP (voir fig. 2) des perpendiculaires dont nous désignerons les pieds respectivement par J'_1, J'_2, B' et T'. Les triangles rectangles CJ'_1J_1 et $BB'C$ d'une part et PJ'_2J_2 et $BB'P$ de l'autre ayant leurs hypothénuses et un angle aigu égaux sont respectivement égaux. Ce qui entraîne : $J_1J'_1 = CB'$, $CJ'_1 = BB'$ et $J_2J'_2 = PB'$, $PJ'_2 = BB'$. Le point T étant

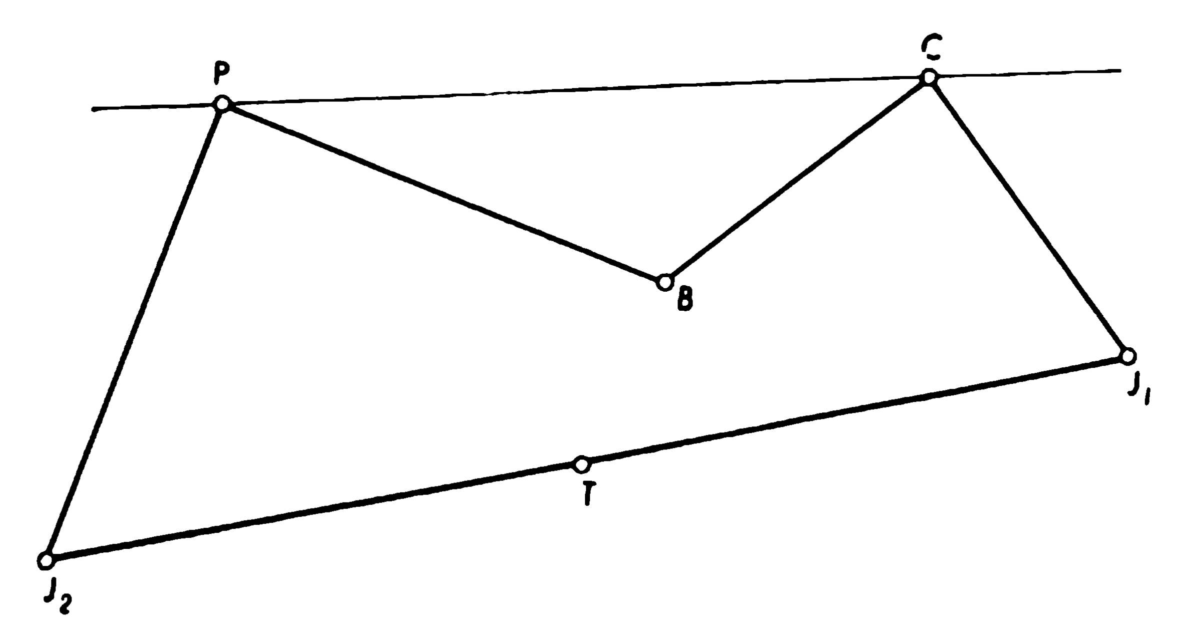


Fig. 1

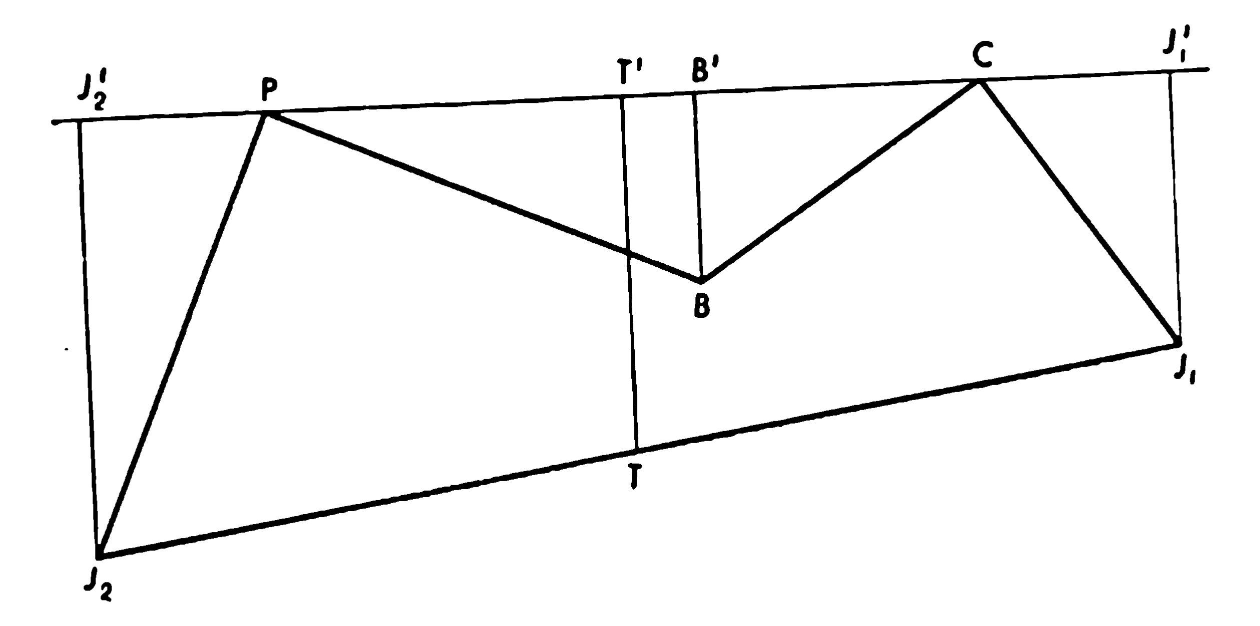


Fig. 2

le milieu du segment J_1J_2, le segment TT' est la ligne médiane du trapèze $J_1'J_1J_2J_2'$, d'où :

$$TT' = \frac{1}{2}(J_1J_1' + J_2J_2') = \frac{1}{2}(CB' + B'P) = \frac{1}{2}CP.$$

D'autre part, puisque T' est le milieu du segment $J'_1J'_2$ et que $CJ'_1 = PJ'_2$ $(= BB')$, on en déduit que T' est le milieu de CP. Donc la position du point T ne dépend pas de celle du point B. Pour trouver le point T, il suffit de tracer la médiatrice du segment CP et de porter sur elle une longueur égale à 1/2 CP dans la direction qui laisse le point C à droite et le point P à gauche.

Bien que la solution proposée soit rigoureuse, elle laisse tout de même planer un sentiment d'insatisfaction. L'idée maîtresse qui a consisté à abaisser à partir des points J_1, J_2, B et T des perpendiculaires sur la droite CP est un pur artifice que l'énoncé du problème ne laisse nullement entrevoir. Il semble en effet plus logique d'établir comment sont liés les points T et B ou plutôt comment varie T en fonction de B. Cette marche est d'ailleurs dictée par la nature même du problème. On s'imagine aisément notre chercheur de trésor dépité de ne point voir le bouleau, parcourir les lieux à la recherche de quelques éventuels vestiges et se tenant le raisonnement suivant: « Si le bouleau s'était trouvé ici, le trésor se serait trouvé là, et si le bouleau s'était trouvé là... ». Et de constater que la cachette du trésor est sans lien avec l'emplacement du bouleau. Il se serait alors emparé de la pelle remettant la démonstration à des moments de meilleure inspiration. Contrairement à ce chercheur, ce qui nous intéresse, nous, c'est la démonstration et non pas seulement la constatation du fait que le point T (i. e. le trésor) ne dépend pas de la position du point B (i. e. du bouleau).

Supposons que le point B commence à se déplacer. Soit $\mathbf{v}$ le vecteur de sa vitesse instantanée. Comme le segment CJ_1 se déduit du segment CB par une rotation d'angle $\frac{\pi}{2}$, le déplacement du point J_1 sera lié à celui du point B: plus exactement le vecteur $\mathbf{v}_1$ de la vitesse instantanée du point J_1 se déduira du vecteur $\mathbf{v}$ par une rotation d'angle $\frac{\pi}{2}$.

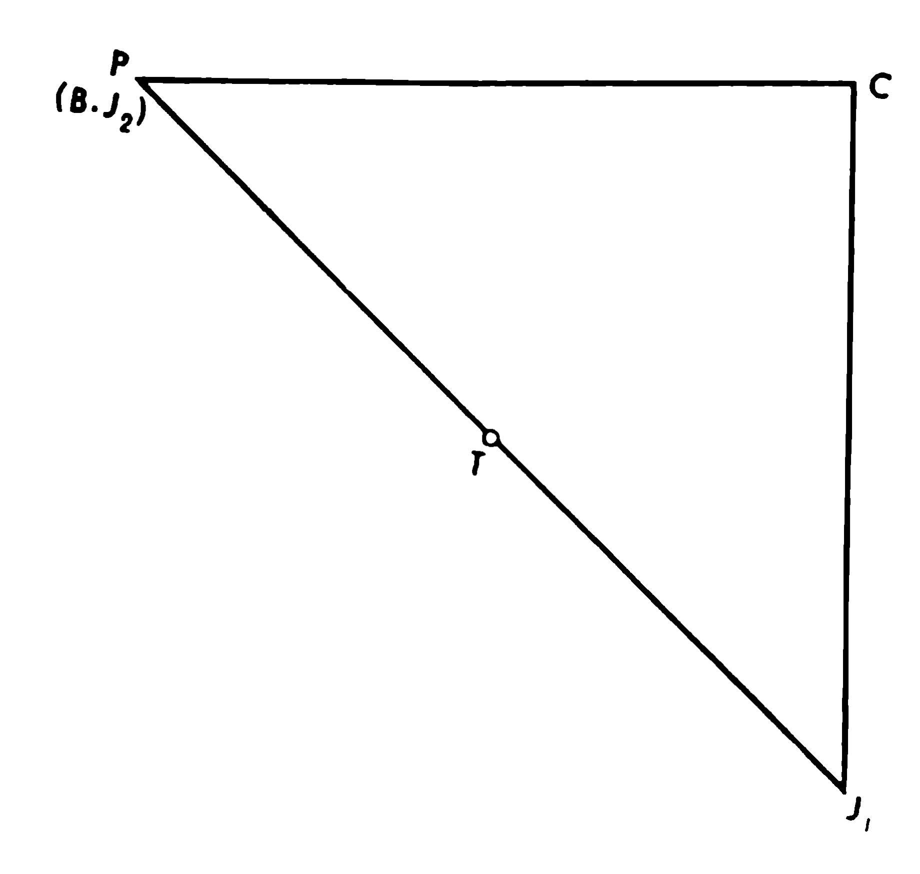


Fig. 3

D'une façon analogue, le vecteur $\mathbf{v}_2$ de la vitesse du point J_2 se déduira de $\mathbf{v}$ par une rotation [1]) d'angle $-\frac{\pi}{2}$. D'où $\mathbf{v}_2 = -\mathbf{v}_1$, et le point T, milieu du segment J_1J_2, se déplacera à la vitesse

$$\mathbf{u} = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2) = 0.$$

Or, si la vitesse d'un point est constamment nulle, c'est que ce point est fixe! Donc, lorsque le point B se déplace arbitrairement, le point T reste fixe. Il en résulte que la position du point T ne dépend pas de celle du point B.

Pour déterminer maintenant la position du point T il suffit de choisir une position quelconque du point B. Le plus simple est de mettre en coïncidence les points B et P et d'utiliser le plan du chercheur du trésor (fig. 3).

---

[1]) Rappelons qu'un angle de rotation est par définition positif ou négatif selon que la rotation a lieu dans le sens ou dans le sens contraire des aiguilles d'une montre.

Cette solution qui procède de notions cinétiques peut en dépit de sa simplicité être inaccessible au lycéen dont les connaissances en mathématiques sont insuffisantes sur les propriétés des vecteurs et des vitesses.

Aussi cette brochure qui traite de l'application des méthodes de cinématique à la résolution de problèmes de géométrie réserve-t-elle une place importante aux vecteurs et aux vitesses. Ces notions étant indispensables à la compréhension de nombreuses branches des mathématiques et de la physique, il importe de les connaître à toute fin utile.

Aux §§ 1 et 2 bien des choses ne sont pas démontrées, mais simplement expliquées. Toutefois en s'inspirant des dessins, le lecteur pourra sans peine trouver une démonstration personnelle. Les lecteurs dont les connaissances sont suffisantes en ce domaine peuvent se contenter d'une rapide lecture de ces deux paragraphes.

Le paragraphe 3 constitue l'ossature de cette brochure. Il développe la solution d'un certain nombre de problèmes par des notions de cinématique et renferme également des exercices.

§ 1 ÉLÉMENTS D'ALGÈBRE VECTORIELLE

nº 1. On appelle *vecteurs* des segments de droite orientés. Les vecteurs sont représentés par des segments fléchés, la flèche en indiquant le sens (fig. 4). L'origine d'un vecteur est appelée également point d'application de ce vecteur. Un vecteur d'origine A et d'extrémité B est noté $\overrightarrow{AB}$ (et non $\overrightarrow{BA}$! qui est le vecteur opposé, i. e. le vecteur d'origine B et d'extrémité A). On désigne souvent un vecteur par une seule lettre par exemple $\overrightarrow{AB} = \mathbf{a}$. En général cette lettre est imprimée en caractère gras pour bien notifier qu'il s'agit d'un vecteur et non d'un nombre. Si un vecteur

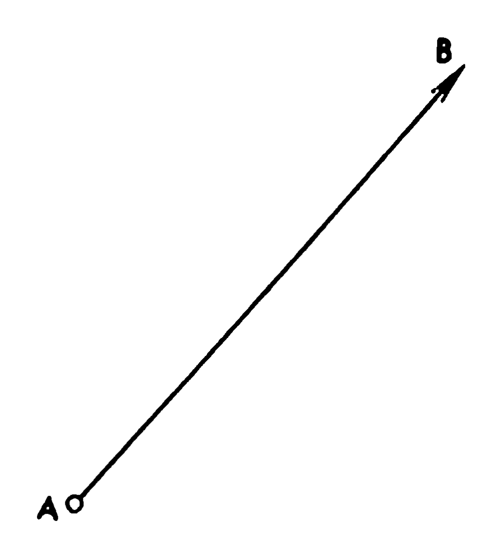


Fig. 4

est désigné par exemple par **a**, sa longueur est notée | **a** |, comme la valeur absolue d'un nombre [1]). La longueur d'un vecteur est souvent appelée *valeur absolue* ou *module*.

L'égalité de deux vecteurs n'est pas comprise au sens d'une identité: elle a un sens plus large. Plus précisément, on dit que deux vecteurs sont *égaux* s'ils ont même module et même sens. Ce qui implique que des vecteurs égaux sont nécessairement parallèles ou ont même support (on dit alors qu'ils sont colinéaires). Sur la fig. 5

$$\overrightarrow{AB} = \overrightarrow{CD}, \quad \overrightarrow{AB} \neq \overrightarrow{PQ}, \quad \overrightarrow{AB} \neq \overrightarrow{EF}, \quad \overrightarrow{AB} \neq \overrightarrow{GH}.$$

De la définition de l'égalité, il résulte qu'un vecteur est invariant dans une translation.

Pour la suite il est important de considérer le point comme un segment dont sont confondues l'origine et l'extrémité. Ce segment « dégénéré » est également un vecteur, mais il n'est doué d'aucun sens [2]). On l'appelle vecteur *nul* et on le note **0**. Son module est égal à zéro: | **0** | = 0.

n°2. On appelle *somme* des vecteurs **a** et **b** le vecteur **c** = **a** + **b** dont l'origine est l'origine du vecteur **a** et

---

[1]) On désigne souvent par $\overline{AB}$ la longueur d'un vecteur $\overrightarrow{AB}$.
[2]) I.e. tout sens est le sens de ce vecteur.

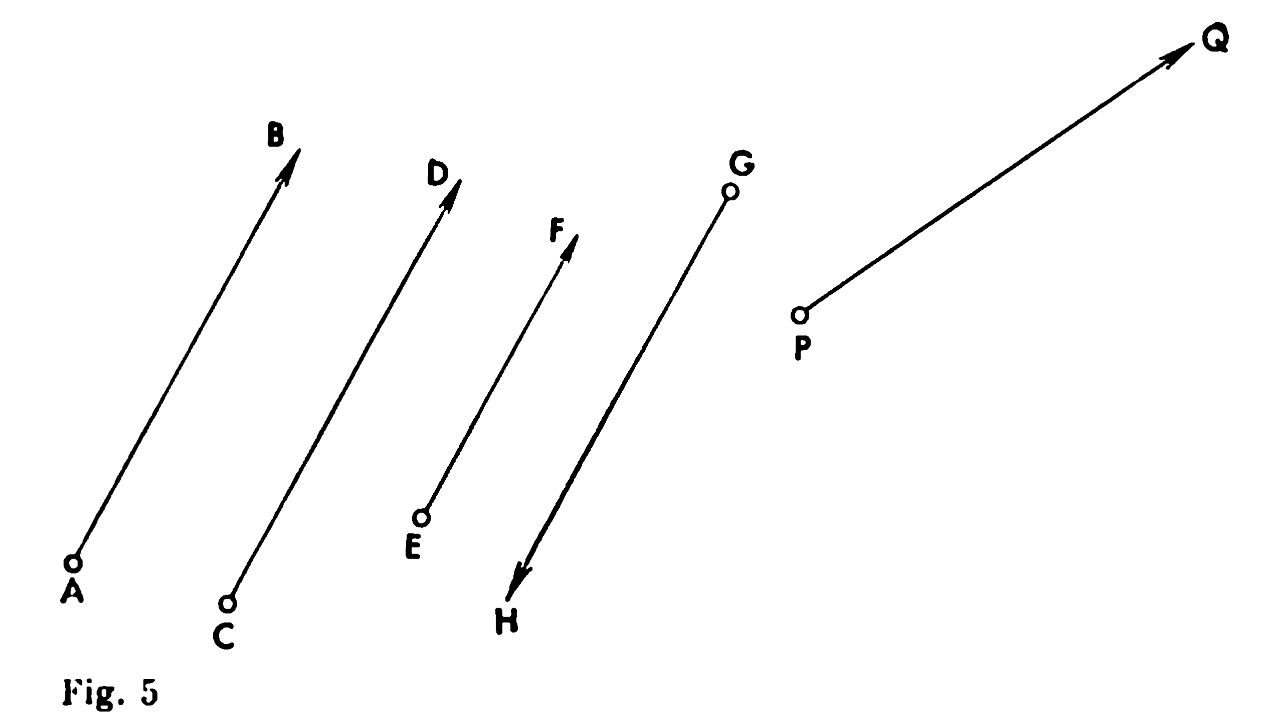


Fig. 5



Fig. 6

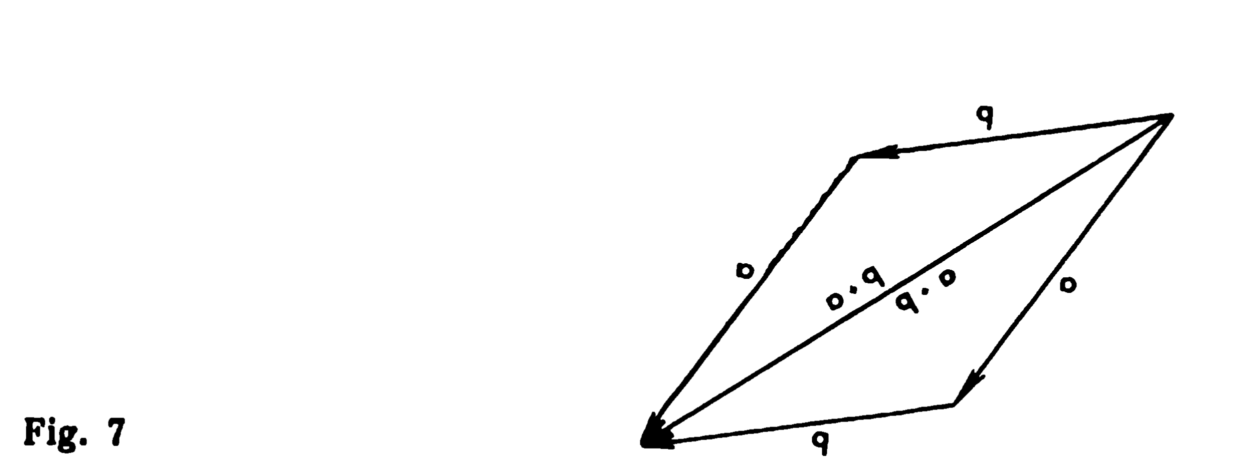


Fig. 7

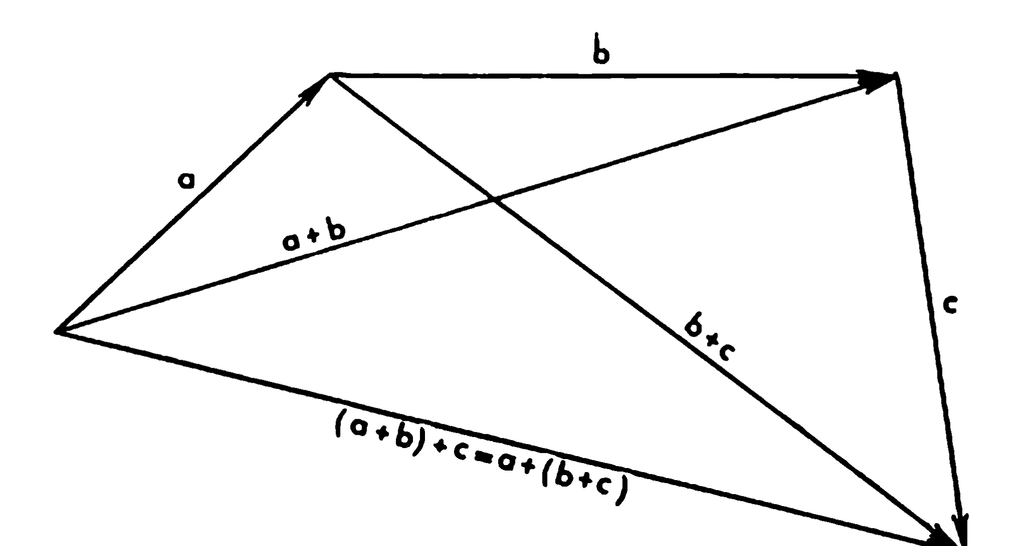


Fig. 8

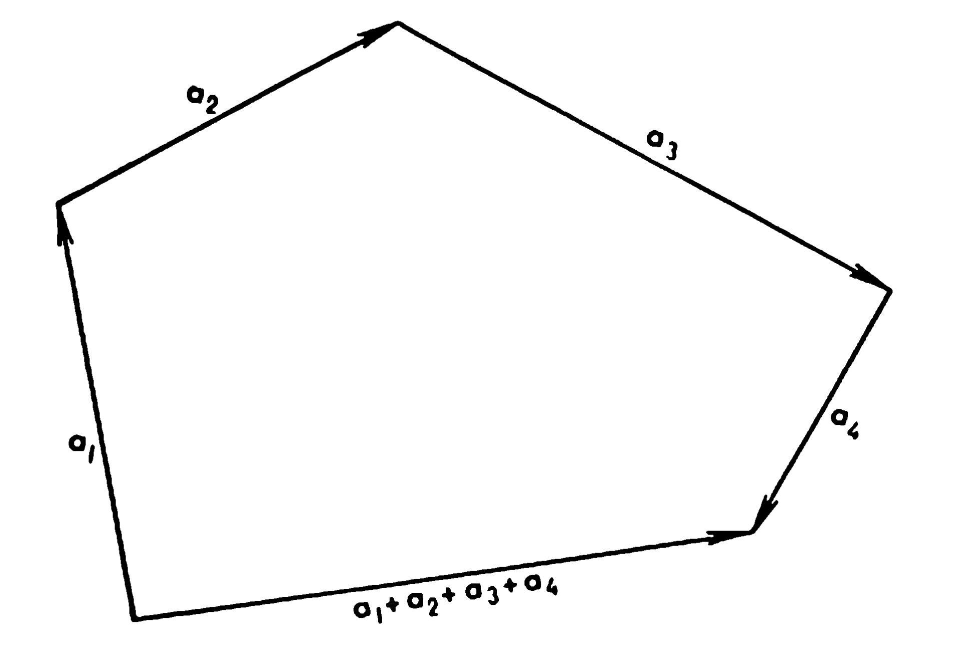


Fig. 9

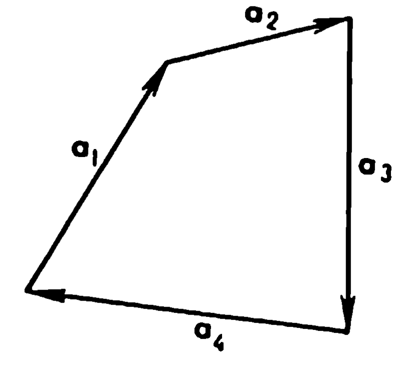


Fig. 10

l'extrémité, l'extrémité du vecteur **b** (voir fig. 6) sous condition que l'origine du vecteur **b** coïncide avec l'extrémité du vecteur **a** (ce que l'on peut réaliser par une translation du vecteur **b**).

La somme des vecteurs comme la somme des nombres est commutative et associative.

La *commutativité* s'exprime par la formule

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \qquad (1)$$

On peut vérifier cette relation sur la fig. 7 où les vecteurs **a** et **b** sont appliqués au même point et sont les côtés d'un parallélogramme dont la diagonale est égale, d'une part à $\mathbf{a} + \mathbf{b}$, et de l'autre, à $\mathbf{b} + \mathbf{a}$.

L'*associativité* est exprimée par la formule

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}). \qquad (2)$$

On vérifie qu'elle est vraie sur la fig. 8.

La commutativité et l'associativité nous permettent comme dans le cas des nombres d'additionner des vecteurs sans tenir compte de l'ordre ni des groupes de termes. En particulier, on peut abaisser les parenthèses et noter tout simplement $\mathbf{a} + \mathbf{b} + \mathbf{c}$.

On obtient la somme de plusieurs vecteurs $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ en les portant bout à bout (voir fig. 9): ils forment une ligne polygonale « bouclée » par le vecteur somme $\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4$.

Il est évident qu'une somme de vecteurs est nulle si et seulement si la ligne polygonale est fermée, i. e. l'extrémité du dernier vecteur coïncide avec l'origine du premier (voir fig. 10).

Supposons par exemple que

$$\mathbf{a} + \mathbf{b} = 0.$$

Les vecteurs **a** et **b** ont donc même module, mais des sens contraires. Le vecteur **b** ainsi défini est appelé *opposé* du vecteur **a** et se note $(-\mathbf{a})$.

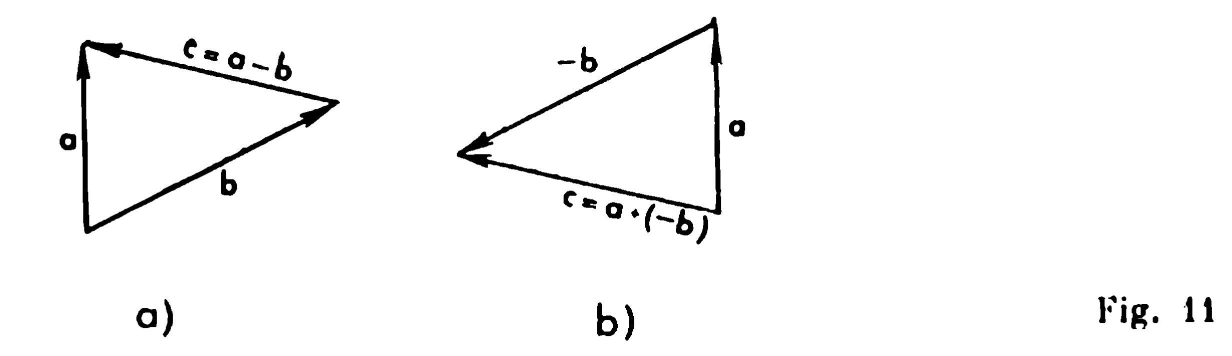


Fig. 11

Les relations

$$\mathbf{a} + 0 = \mathbf{a}, \qquad \mathbf{a} + (-\mathbf{a}) = 0, \tag{3}$$

qui découlent immédiatement de la définition, sont importantes en algèbre vectorielle. Elles permettent en particulier d'étudier la différence des vecteurs qui est l'inverse de la somme.

n°3. On appelle *différence* $\mathbf{a} - \mathbf{b}$ des vecteurs $\mathbf{a}$ et $\mathbf{b}$ le vecteur $\mathbf{c}$ tel que

$$\mathbf{b} + \mathbf{c} = \mathbf{a}. \tag{4}$$

La méthode de construction de la différence est indiquée sur la fig. 11, *a*. On peut néanmoins ramener la différence à une somme en procédant comme suit. Ajoutons aux deux membres de l'égalité (4) le vecteur $(-\mathbf{b})$:

$$\mathbf{a} + (-\mathbf{b}) = (\mathbf{b} + \mathbf{c}) + (-\mathbf{b}).$$

En vertu de l'associativité et de la commutativité on a :

$$\mathbf{a} + (-\mathbf{b}) = \mathbf{c} + [\mathbf{b} + (-\mathbf{b})],$$

d'où, compte tenu de (3),

$$\mathbf{a} + (-\mathbf{b}) = \mathbf{c} + 0 = \mathbf{c}.$$

Donc

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}). \tag{5}$$

Ce qui nous fournit un autre procédé de construction de la différence (voir fig. 11, *b*).

Notons encore les formules:

$$a - 0 = a, \quad 0 - a = -a, \quad a - a = 0. \qquad (6)$$

Puisqu'en vertu de la définition de la différence, les égalités

$$a - b = c, \qquad a = b + c$$

ont la même signification, on peut faire passer un vecteur d'un membre à l'autre à condition de lui changer son signe.

n°4. Nous aurons besoin d'une importante inégalité appelée *inégalité du triangle*. Considérons la fig. 6 (page 221). Entre les trois côtés d'un triangle nous avons la relation

$$|a + b| \leqslant |a| + |b|. \qquad (7)$$

L'égalité est réalisée si et seulement si les vecteurs ont le même sens.

Signalons encore d'autres inégalités analogues à celle du triangle

$$|a - b| \leqslant |a| + |b|; \qquad |a - b| \geqslant |a| - |b|. \quad (8)$$

n°5. On appelle *produit* λa du vecteur a par le nombre réel λ [1]) le vecteur [2]) c tel que:

1) $|c| = |\lambda| \cdot |a|$, où $|\lambda|$ est la valeur absolue du nombre λ;

2) c et a sont colinéaires;

3) le vecteur c est de même sens que le vecteur a lorsque $\lambda > 0$, et de sens contraire lorsque $\lambda < 0$. Sur la fig. 12, $\lambda = 3/2$ et $\lambda = -3/4$. Il est évident que

$$a = 1 \cdot a, \quad a + a = 2a, \quad a + a + a = 3a, \ldots$$

et

$$(-a) = (-1) \cdot a, \quad (-a) + (-a) = (-2) \cdot a;$$

$$(-a) + (-a) + (-a) = (-3) \cdot a, \ldots$$

---

[1]) Le nombre λ est appelé aussi scalaire.

[2]) On obtient un vecteur et non un nombre.

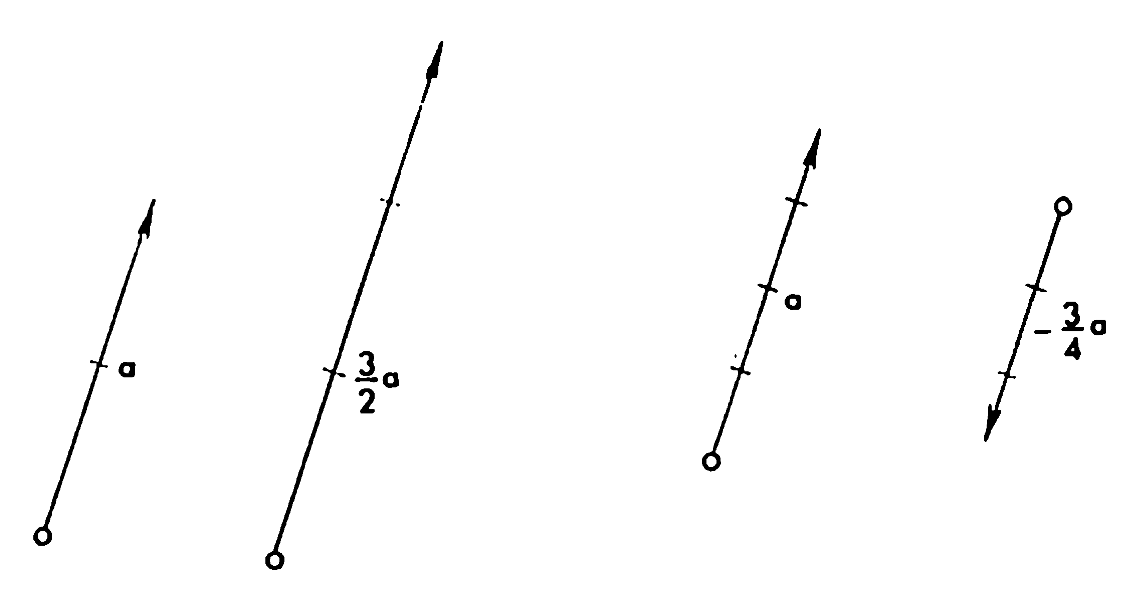


Fig. 12

Indiquons les principales propriétés du produit d'un vecteur par un scalaire.

1) *Le produit est associatif*

$$\mu(\lambda \mathbf{a}) = (\mu\lambda)\,\mathbf{a}. \qquad (9)$$

Sur la fig. 13 sont représentés les cas où $\lambda > 0$, $\mu > 0$ et $\lambda > 0$, $\mu < 0$.

2) *Le produit est distributif par rapport à l'addition des vecteurs*

$$\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}. \qquad (10)$$

Sur la fig. 14 sont représentés les cas où $\lambda > 0$ et $\lambda < 0$.

3) *Le produit est distributif par rapport à l'addition des scalaires*

$$(\lambda + \mu)\,\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}. \qquad (11)$$

Sur la fig. 15 sont représentés les cas où $\lambda > 0$, $\mu > 0$ et $\lambda > 0$, $\mu < 0$, $\lambda + \mu > 0$. Signalons encore une égalité évidente

$$0\cdot\mathbf{a} = 0, \quad \lambda\cdot 0 = 0. \qquad (12)$$

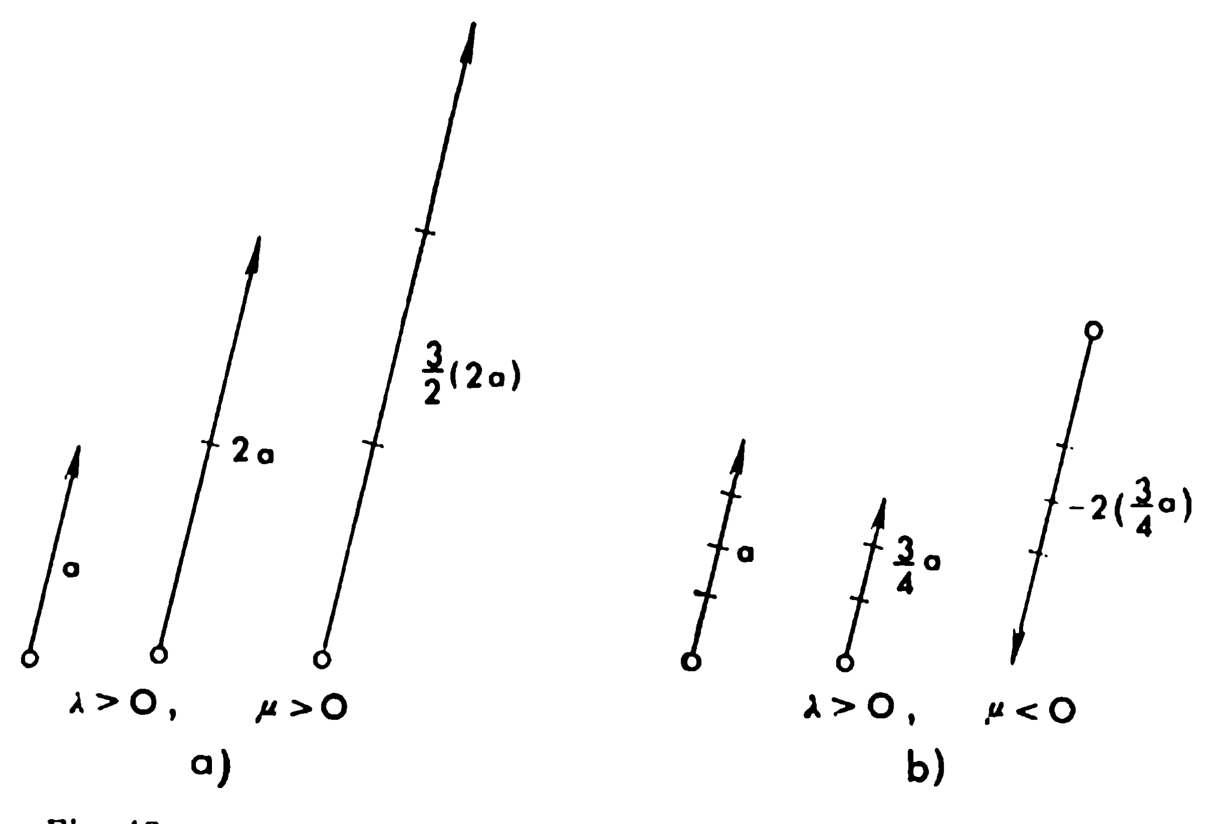

a
2a
$\frac{3}{2}(2a)$
$\lambda > 0$, $\mu > 0$
a)
a
$\frac{3}{4}a$
$-2(\frac{3}{4}a)$
$\lambda > 0$, $\mu < 0$
b)

Fig. 13

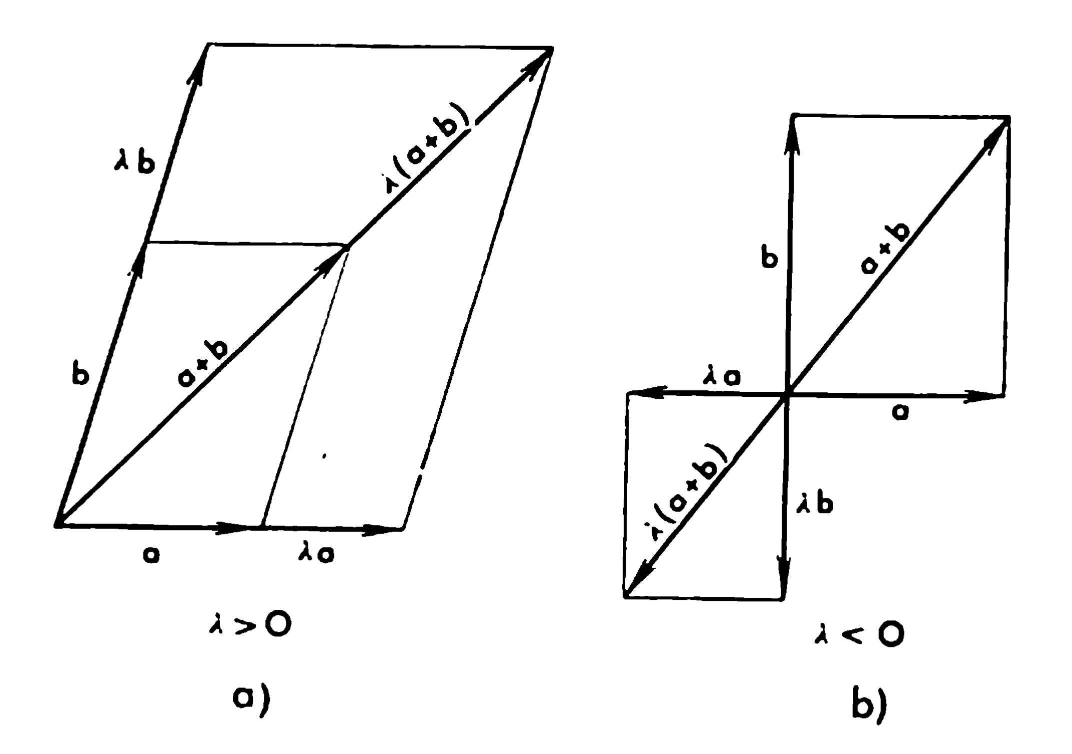

λb
$\lambda(a+b)$
b
a+b
a
λa
$\lambda > 0$
a)
b
a+b
λa
a
$\lambda(a+b)$
λb
$\lambda < 0$
b)

Fig. 14

On peut également considérer la *division d'un vecteur par un scalaire.* On appelle division du vecteur **a** par le scalaire $\lambda \neq 0$, le produit du vecteur **a** par l'inverse de ce scalaire:

$$\frac{\mathbf{a}}{\lambda} = \frac{1}{\lambda} \cdot \mathbf{a}. \quad (13)$$

Nous remarquons donc que les lois fondamentales de l'algèbre vectorielle ne diffèrent pas de celles de l'algèbre numérique. C'est pourquoi en algèbre vectorielle les opérations s'effectuent comme en algèbre numérique.

n°6. Dans toute la suite nous supposerons que les vecteurs sont contenus dans un même plan.

Soient **a** et **b** deux vecteurs non colinéaires et **c** un vecteur quelconque. Si le vecteur **c** est colinéaire avec l'un des vecteurs **a** ou **b**, disons **a**, il existe un scalaire λ tel que

$$\mathbf{c} = \lambda \mathbf{a}. \quad (14)$$

Dans le cas général, appliquons les trois vecteurs en un point O (fig. 16). A partir de l'extrémité C du vecteur **c** menons des parallèles aux vecteurs **a** et **b**. Soient A et B les points d'intersection de ces parallèles avec les supports de **a** et **b**.

Il est évident que

$$\mathbf{c} = \overrightarrow{OA} + \overrightarrow{OB}.$$

Or, comme les vecteurs $\overrightarrow{OA}$ et **a** sont colinéaires, il existe un λ tel que

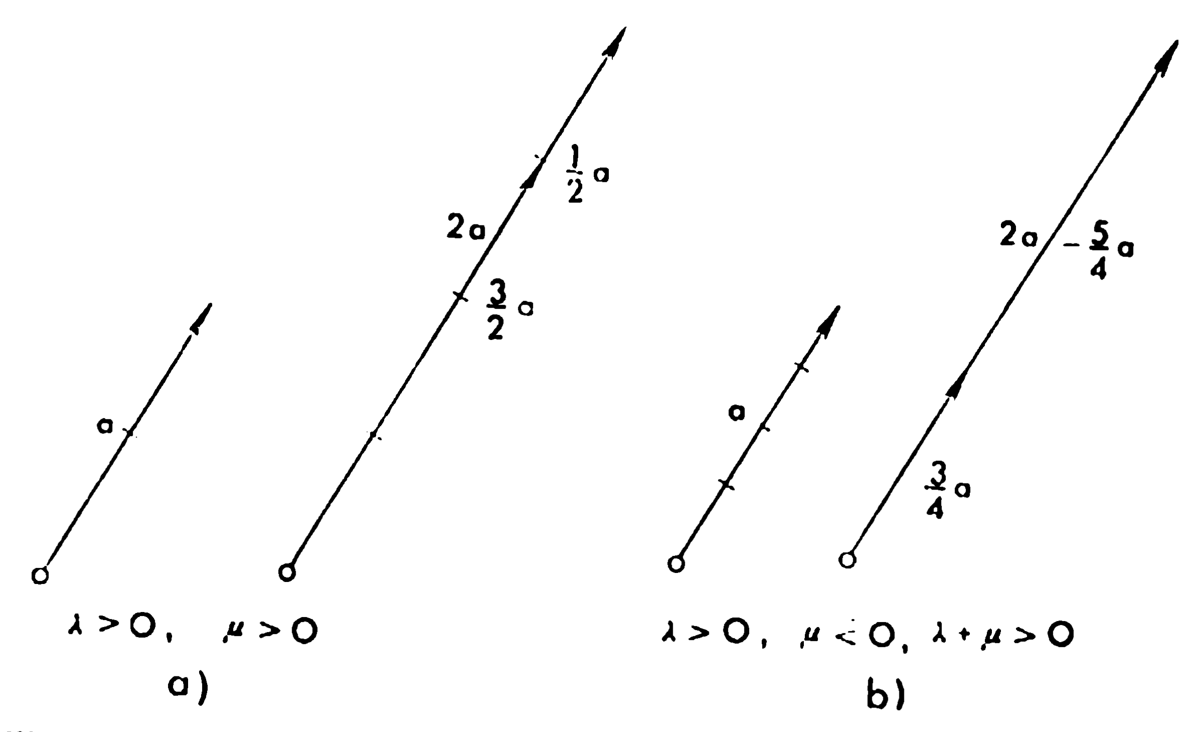
$$\overrightarrow{OA} = \lambda \mathbf{a}.$$

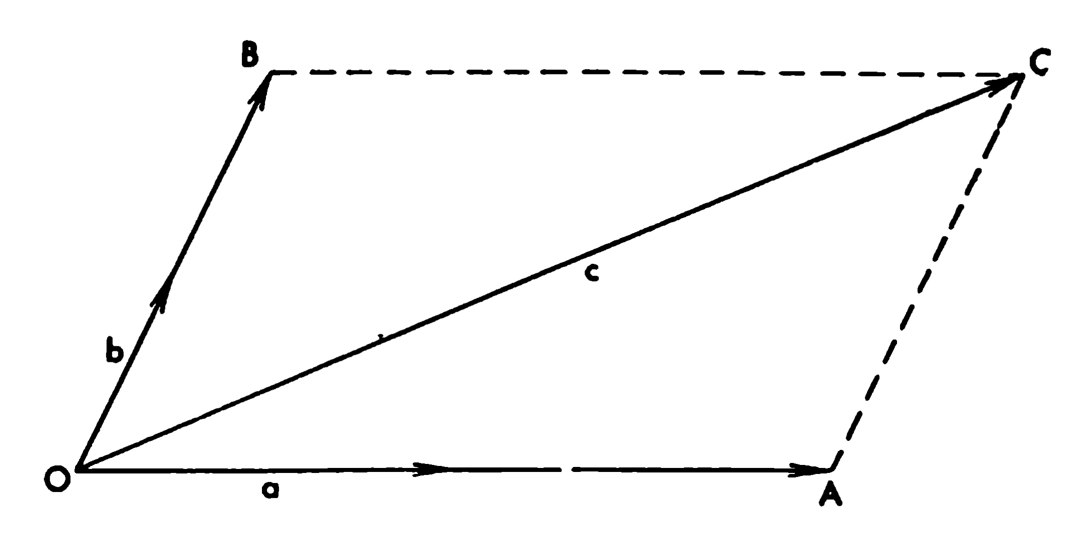
D'une façon analogue, il existe un nombre μ tel que

$$\overrightarrow{OB} = \mu \mathbf{b}.$$

Donc:

$$\mathbf{c} = \lambda \mathbf{a} + \mu \mathbf{b}. \quad (15)$$


a
2a
$\frac{1}{2}$ a
$\frac{3}{2}$ a
λ > O, μ > O
a)
a
2a
$-\frac{5}{4}$ a
$\frac{3}{4}$ a
λ > O, μ < O, λ + μ > O
b)
Fig. 15


B
C
c
b
O
a
A
Fig. 16

La représentation du vecteur **c** sous la forme (15) s'appelle *décomposition* de ce vecteur suivant les vecteurs **a** et **b**. Tout vecteur **c** est susceptible d'être décomposé suivant deux vecteurs **a** et **b** non colinéaires, les coefficients λ et μ étant définis d'une manière univoque.

Observons que l'égalité (14) s'écrit sous la forme (15) avec $\mu = 0$.

n°7. Soient A, B et C trois points quelconques alignés. On dit que le point C divise le segment AB dans le rapport $m : n$ si [1])

$$n\overrightarrow{AC} = m\,\overrightarrow{CB}. \qquad (16)$$

Il est évident que la valeur absolue du rapport $m : n$ est égale au rapport des longueurs $AC : CB$. Le rapport $m : n$ est positif si le point C est contenu entre A et B et négatif si C est situé hors du segment AB (fig. 17).

● THÉORÈME. *Supposons que le point C divise le segment AB dans le rapport $m : n$ et soit O un point quelconque du plan* (fig. 18). *Alors*

$$\overrightarrow{OC} = \frac{n\overrightarrow{OA} + m\overrightarrow{OB}}{m+n}. \qquad (17)$$

Réciproque. *Si la relation* (17) *est vérifiée pour un point quelconque O, alors le point C divise le segment AB dans le rapport $m : n$.*

● DÉMONSTRATION. Supposons la relation (16) vérifiée. Comme

$$\overrightarrow{AC} = \overrightarrow{OC} - \overrightarrow{OA}, \qquad \overrightarrow{CB} = \overrightarrow{OB} - \overrightarrow{OC},$$

donc

$$n(\overrightarrow{OC} - \overrightarrow{OA}) = m(\overrightarrow{OB} - \overrightarrow{OC}).$$

---

[1]) m et n sont des nombres réels quelconques, non simultanément nuls. Si $m = 0$, le point C se confond avec A ; si $n = 0$, il se confond avec B.

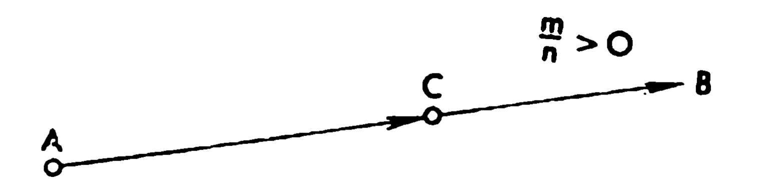




Fig. 17

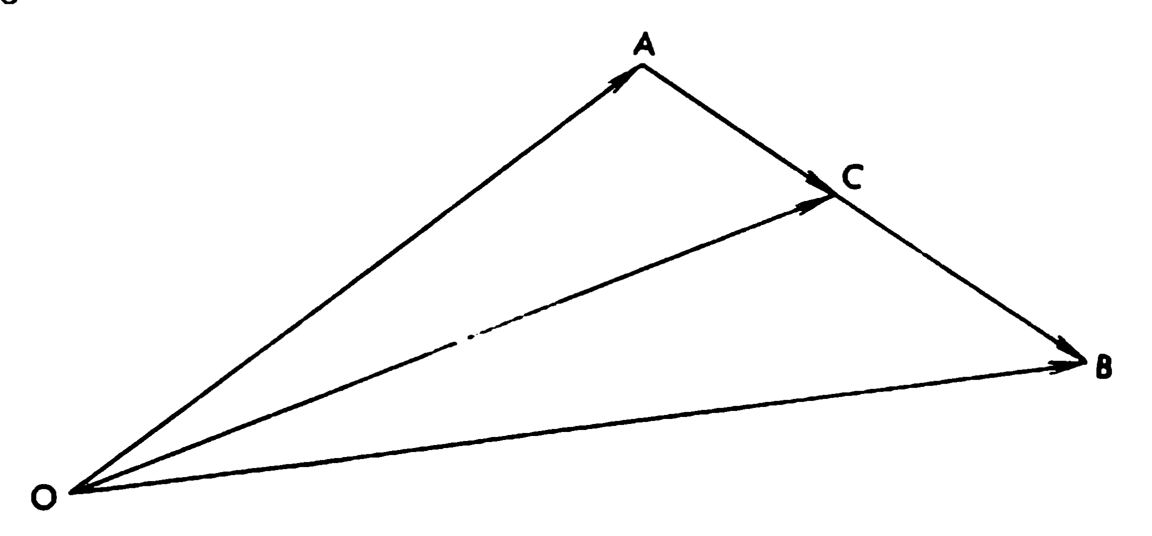


Fig. 18

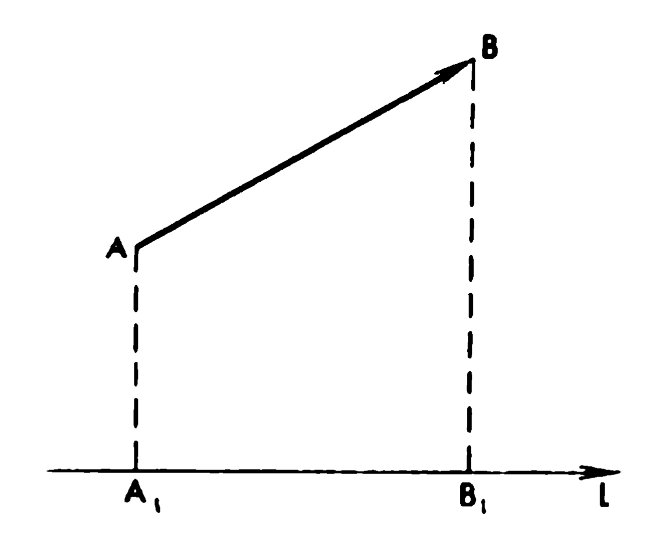


Fig. 19

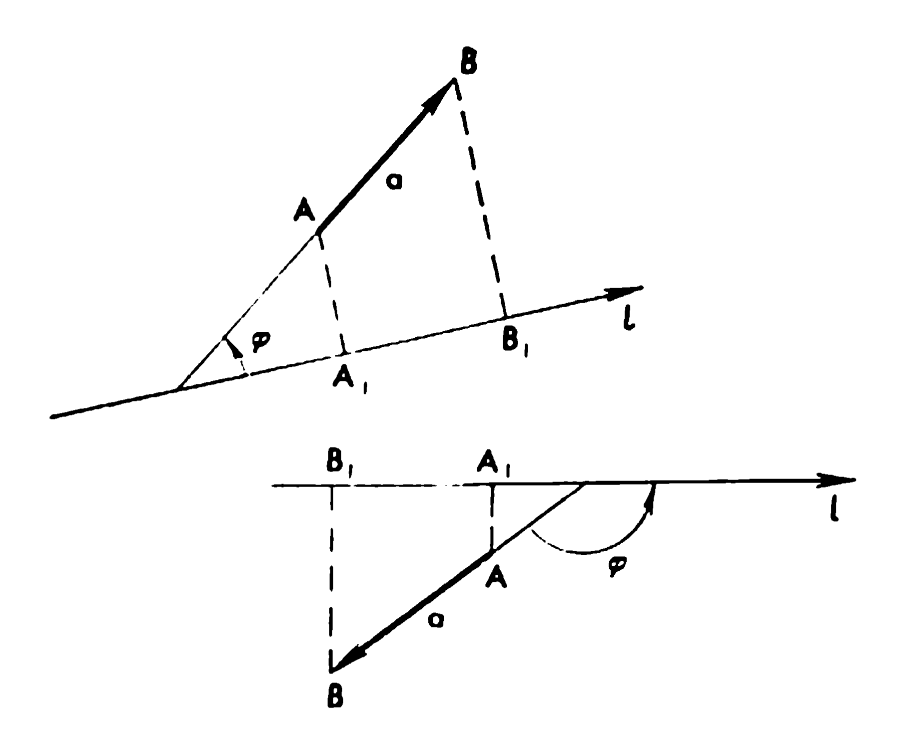


Fig. 20

Si l'on résout cette équation relativement à $\overrightarrow{OC}$, on obtient (17).

D'une façon analogue, de (17) on obtient (16).

n°8. On sait qu'un *axe* est une droite sur laquelle on a défini un sens de parcours.

Soient un axe l et un vecteur arbitraire $\overrightarrow{AB}$ (fig. 19). Désignons par A_l et B_l les projections orthogonales des points A et B sur l'axe l (i. e. les pieds des perpendiculaires abaissées de A et B sur l). Considérons le nombre égal à la longueur du segment A_lB_l précédé du signe plus si le sens du vecteur $\overrightarrow{A_lB_l}$ est celui de l'axe, et du signe moins dans le cas contraire. Ce nombre est appelé mesure algébrique de la projection du vecteur $\overrightarrow{AB}$ sur l'axe l, on le note $\text{proj}_l\,\overrightarrow{AB}$.

Désignons par φ l'angle que fait le vecteur **a** avec l'axe l (fig. 20). L'angle φ est compris entre 0 et π. Il est évident que

$$\text{proj}_l\,\mathbf{a} = |\,\mathbf{a}\,|\cdot\cos\varphi. \tag{18}$$

En particulier, si $\mathbf{a}$ est perpendiculaire à l, on a $\text{proj}_l\,\mathbf{a} = 0$.

Notons encore deux propriétés de la projection (fig. 21, 22):

1) $\text{proj}_l\,(\mathbf{a} + \mathbf{b}) = \text{proj}_l\,\mathbf{a} + \text{proj}_l\,\mathbf{b}$.

2) $\text{proj}_l\,(\lambda\mathbf{a}) = \lambda\,\text{proj}_l\,\mathbf{a}$ (λ est un scalaire quelconque).

Ces propriétés s'énoncent comme suit: « la projection d'un vecteur sur un axe est une *opération linéaire* sur les vecteurs ». En utilisant les propriétés 1) et 2) on obtient en général:

$$\text{proj}_l\,(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_n\mathbf{a}_n) = \\ = \lambda_1\,\text{proj}_l\mathbf{a}_1 + \lambda_2\,\text{proj}_l\mathbf{a}_2 + \ldots + \lambda_n\,\text{proj}_l\mathbf{a}_n \quad (19)$$

quels que soient les vecteurs $\mathbf{a}_1$, $\mathbf{a}_2$, . . ., $\mathbf{a}_n$ et les scalaires λ_1, λ_2, . . ., λ_n.

Soulignons que le produit d'un vecteur par un scalaire λ est aussi une opération linéaire (voir (9), (10)).

n°9. Il existe encore un exemple important d'opération linéaire, c'est la rotation d'un vecteur d'un angle α donné, positif, négatif ou nul. Désignons cette opération par U_α. Ainsi le vecteur $U_\alpha\mathbf{a}$ se déduit du vecteur $\mathbf{a}$ par une rotation d'angle α, et de plus

$$|\,U_\alpha\mathbf{a}\,| = |\,\mathbf{a}\,| \quad (20)$$

(fig. 23).

Il est évident que $U_0\mathbf{a} = \mathbf{a}$, i. e. l'opération U_0 laisse invariant le vecteur $\mathbf{a}$. Toute opération qui laisse invariant un vecteur est appelée *opération identique*.

Remarquons encore que

$$U_\pi\mathbf{a} = -\mathbf{a}, \qquad U_{2\pi}\mathbf{a} = \mathbf{a}. \quad (21)$$

Comme l'opération U_α est linéaire:

1) $U_\alpha\,(\mathbf{a} + \mathbf{b}) = U_\alpha\mathbf{a} + U_\alpha\mathbf{b}$ (fig. 24),

2) $U_\alpha\,(\lambda\mathbf{a}) = \lambda U_\alpha\mathbf{a}$, où λ est un scalaire quelconque (fig. 25). Donc, par analogie avec (19),

$$U_\alpha\,(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_n\mathbf{a}_n) = \\ = \lambda_1 U_\alpha\mathbf{a}_1 + \lambda_2 U_\alpha\mathbf{a}_2 + \ldots + \lambda_n U_\alpha\mathbf{a}_n. \quad (22)$$

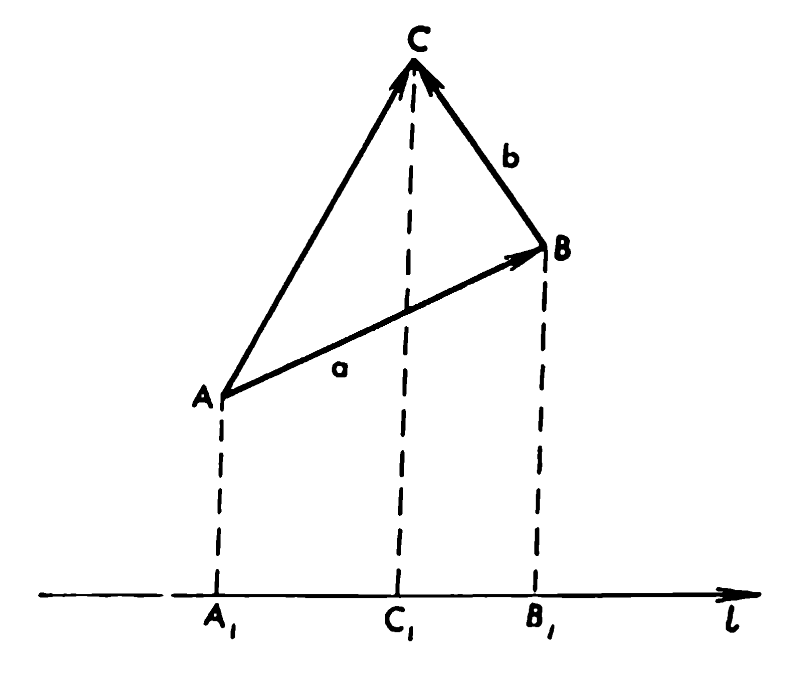


Fig. 21

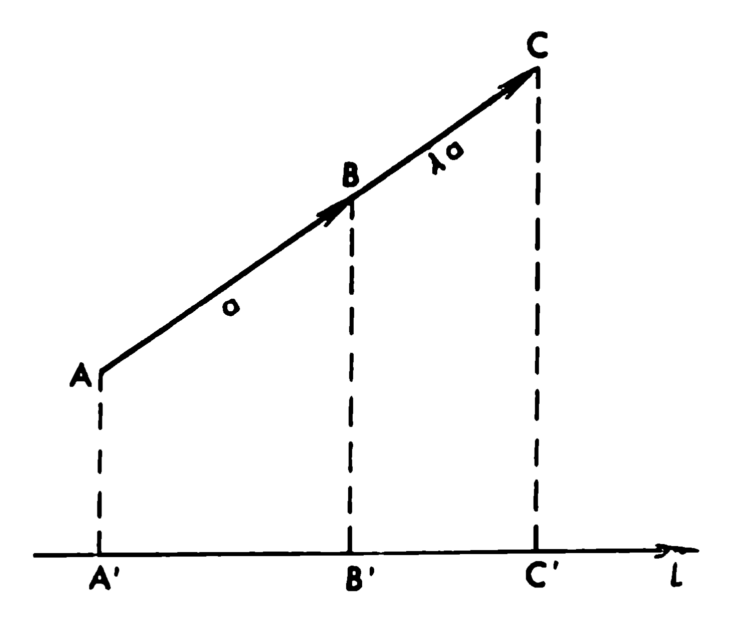
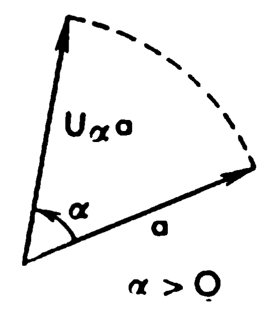


Fig. 22



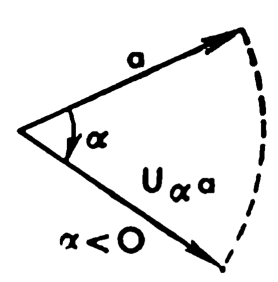


Fig. 23

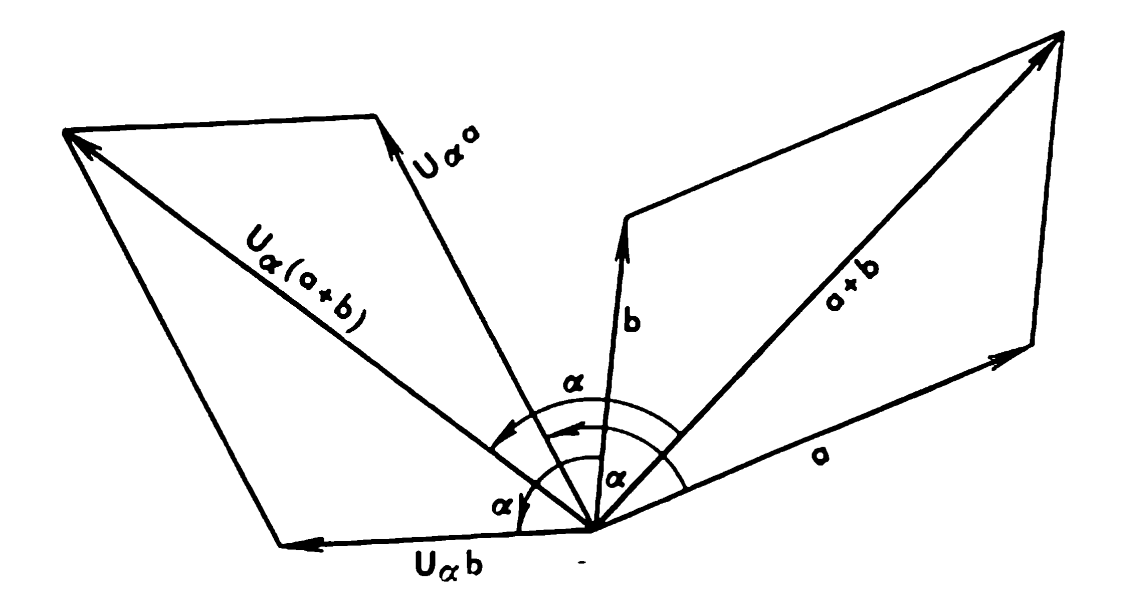


Fig. 24

n°10. Soient S et T deux opérations sur les vecteurs (exemple une projection S sur un axe l et une rotation T d'un angle droit). La réalisation successive de ces opérations est appelée *produit de ces opérations*. Faisons observer que le résultat de ce produit dépend en général de l'ordre dans lequel sont réalisées ces opérations. Si on prend un vecteur $\mathbf{a} \neq 0$ porté par l'axe l et qu'on lui applique d'abord l'opération T, et ensuite l'opération S, on obtient 0; si, au contraire, on lui applique l'opération S, on obtient un nombre auquel il est impossible d'appliquer l'opération T (celle-ci ne s'appliquant qu'à des vecteurs).

Si on réalise d'abord l'opération T, et ensuite l'opération S, on note leur produit ST. Donc, par définition, on a :

$$(ST)\,\mathbf{a} = S\,(T\mathbf{a}) \tag{23}$$

quel que soit $\mathbf{a}$.

La fig. 26 montre que $U_\beta U_\alpha \mathbf{a} = U_{\alpha+\beta}\mathbf{a}$ quel que soit $\mathbf{a}$, i. e.

$$U_\beta U_\alpha = U_{\alpha+\beta}\ ^{1)}. \tag{24}$$

[1]) Deux opérations T_1, T_2 sur les vecteurs sont identiques si $T_1\mathbf{a} = T_2\mathbf{a}$ quel que soit $\mathbf{a}$.

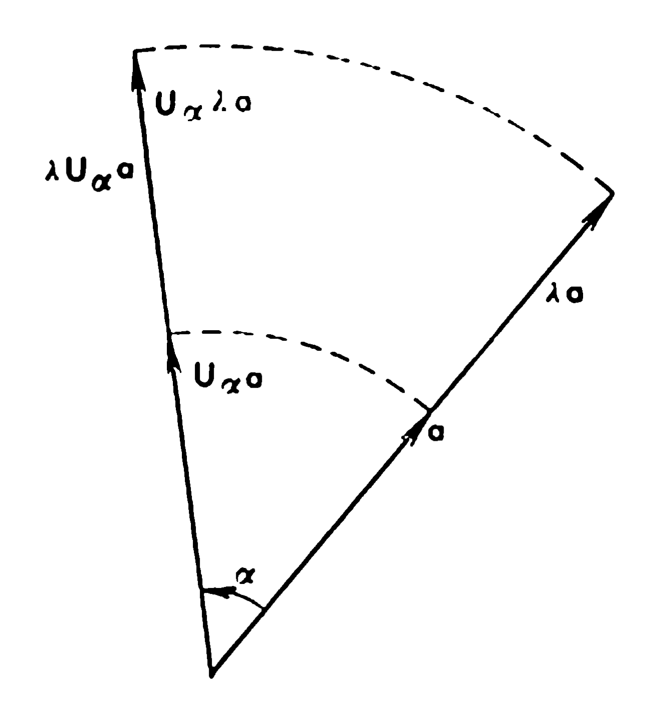


Fig. 25

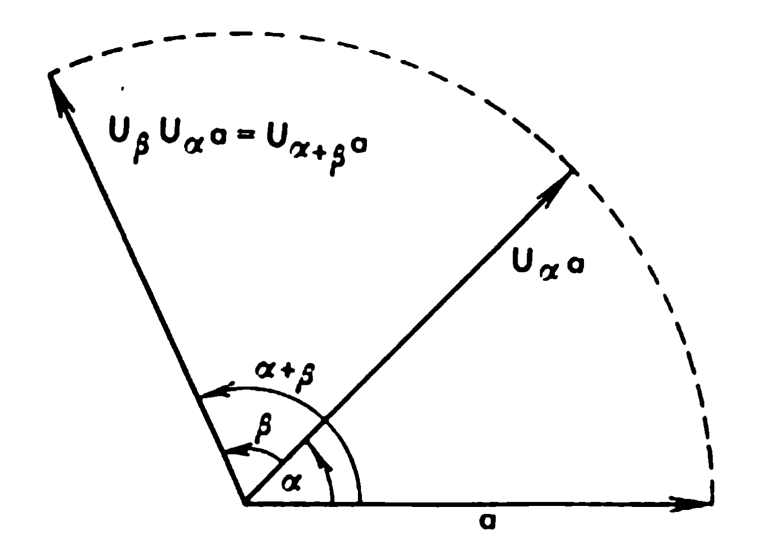


Fig. 26

Il est clair que

$$U_\alpha U_\beta = U_\beta U_\alpha, \tag{25}$$

quoique d'une manière générale $ST \neq TS$.

Si $ST = TS$, on dit que les opérations S et T sont *commutables* ou que leur produit est *commutatif*. Ainsi le produit de deux rotations est commutatif de même que le produit d'une rotation par un nombre (propriété 2 de l'opération U_α).

La formule (24) entraîne:

$$U_{-\alpha} U_\alpha = U_0,$$

i. e. $U_{-\alpha} U_\alpha \mathbf{a} = \mathbf{a}$ quel que soit le vecteur **a**.

Les propriétés des rotations et des autres transformations géométriques peuvent être utilisées avec succès dans la résolution des problèmes les plus variés.

§ 2 ÉLÉMENTS DE CINÉMATIQUE

n°1. Soit sur le plan un point quelconque O que nous appellerons *pôle*. Le vecteur $\mathbf{r} = \overrightarrow{OM}$ (fig. 27) est par définition le *rayon vecteur* du point M relativement au pôle O. Un point et son rayon vecteur se définissent donc mutuellement.

Si un point décrit une trajectoire quelconque (fig. 28), son rayon vecteur varie par rapport au temps: il est donc *fonction* du temps et on le note:

$$\mathbf{r} = \mathbf{r}(t) \quad (t \text{ désigne la variable temps}). \qquad (1)$$

Le mot « varie » ne doit pas être compris à la lettre. Un cas particulier important du mouvement est l'état de repos. Si un point est au repos, son rayon vecteur ne se modifiera pas avec le temps. Il est donc une constante relativement au temps. Ce qu'on note

$$\mathbf{r} = \text{const.} \qquad (2)$$

Lorsqu'on dit par exemple que **r** est une fonction de t, on veut par là exprimer le fait qu'à chaque instant t le vecteur **r** est entièrement défini: si t est fixe, alors **r** est constant.

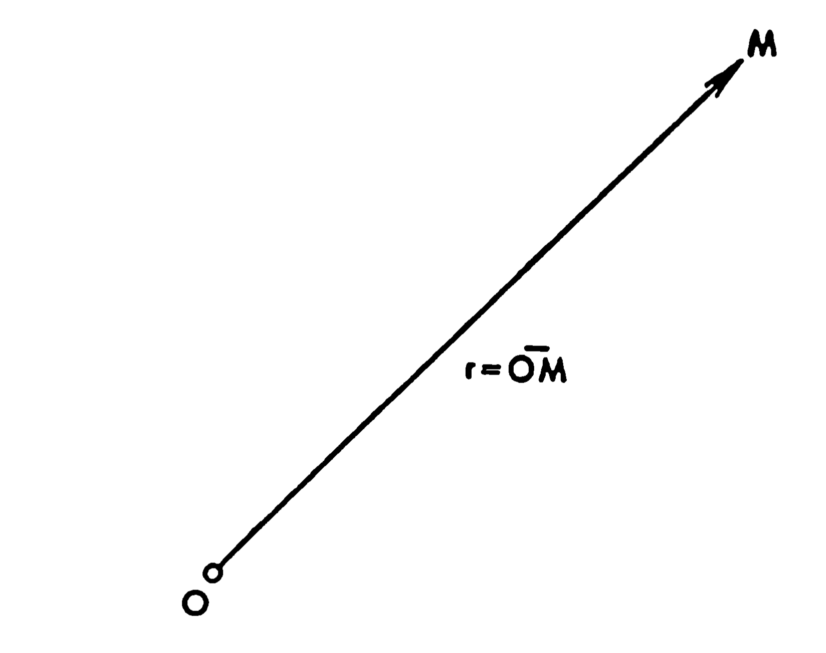


Fig. 27

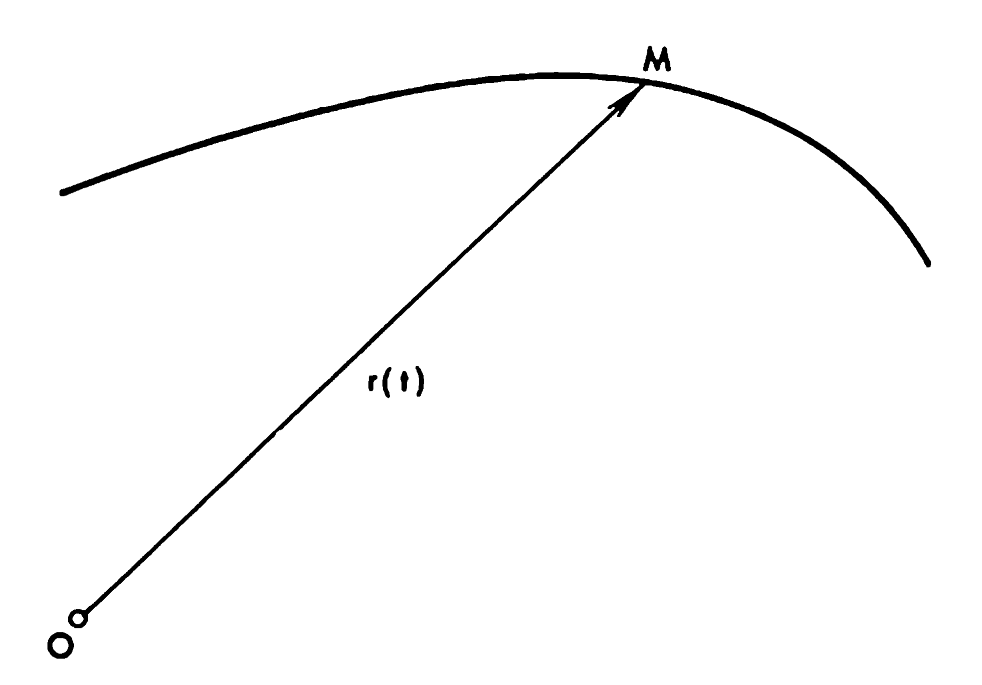


Fig. 28

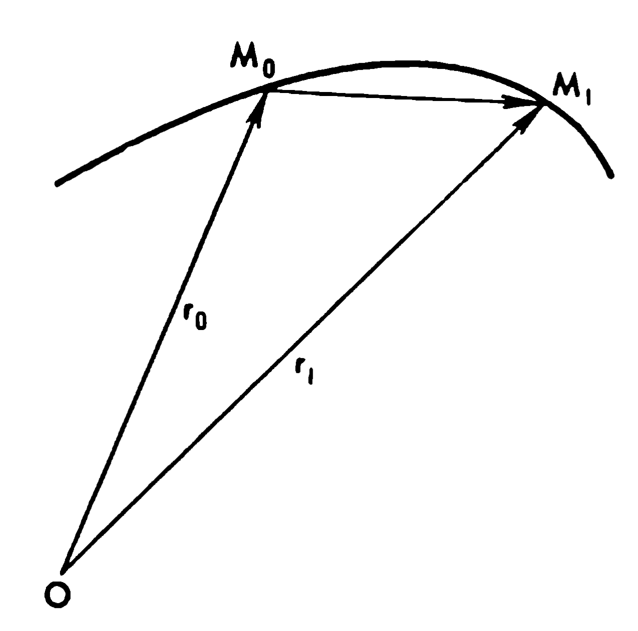


Fig. 29

n°2. Considérons un intervalle quelconque de temps $[t_0, t_1]$. Sa longueur est [1]

$$\Delta t = t_1 - t_0. \tag{3}$$

Si à l'instant t_0 le rayon vecteur du point mobile M est égal à $\mathbf{r}_0$ ($\mathbf{r}_0 = \mathbf{r}(t_0)$), et à l'instant t_1 à $\mathbf{r}_1$ ($\mathbf{r}_1 = \mathbf{r}(t_1)$), alors le vecteur

$$\Delta \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_0$$

représente le déplacement du point M pendant l'intervalle de temps $[t_0, t_1]$ (fig. 29).

Introduisons maintenant une importante notion: celle de *vitesse d'un mouvement*. D'une façon grossière, nous appellerons vitesse le déplacement dans l'unité de temps. Ceci étant, la vitesse doit exprimer le module du déplacement dans l'unité de temps et le sens de ce déplacement, i. e. elle doit être un vecteur.

---

[1]) Le symbole Δ sert d'ordinaire à désigner un accroissement d'une quantité quelconque, i.e. de combien a varié cette quantité.

Si nous savons que pendant l'intervalle de temps $[t_0, t_1]$ le point M a subi un déplacement $\Delta \mathbf{r}$, pour obtenir le déplacement de ce point dans l'unité de temps, il convient de diviser $\Delta \mathbf{r}$ par la longueur de cet intervalle de temps. Ce faisant nous obtenons un vecteur appelé *vitesse moyenne du point M dans l'intervalle de temps donné*, *que l'on note*:

$$\mathbf{v}_m = \frac{\Delta \mathbf{r}}{\Delta t}. \tag{4}$$

Ce vecteur est orienté dans le même sens que le vecteur déplacement $\Delta \mathbf{r}$, mais son module est égal au rapport de la distance M_0M_1 à Δt, i. e. d'une façon grossière, à la distance parcourue dans l'unité de temps.

Qu'entendons-nous par « d'une façon grossière »? Soulignons en effet que le mouvement du point M pendant l'intervalle $[t_0, t_1]$ est en général non uniforme, i. e. le point M parcourt des distances inégales pendant des sous-intervalles de temps égaux. De plus, ce point se déplace d'une façon générale non pas sur une droite passant par M_0 et M_1, mais suivant une courbe reliant ces deux points. Le vecteur $\Delta \mathbf{r}$ caractérise uniquement le bilan de ce mouvement et non pas ses états intermédiaires. Ceci concerne également le vecteur vitesse moyenne, d'où l'épithète « moyenne ».

Cependant, il est aisé de comprendre que la vitesse moyenne caractérisera convenablement le mouvement si l'intervalle de temps est suffisamment petit. Aussi pour décrire idéalement la vitesse importe-t-il de faire tendre Δt vers zéro, i. e., t_0 étant fixe, de faire tendre t_1 vers t_0. La vitesse moyenne $\mathbf{v}_m$ tendra généralement vers une certaine limite $\mathbf{v}$:

$$\mathbf{v} = \lim_{\Delta t \to 0}, \quad \mathbf{v}_{moy} = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{r}}{\Delta t}. \tag{5}$$

Ceci est dans une certaine mesure illustré sur la fig. 30.

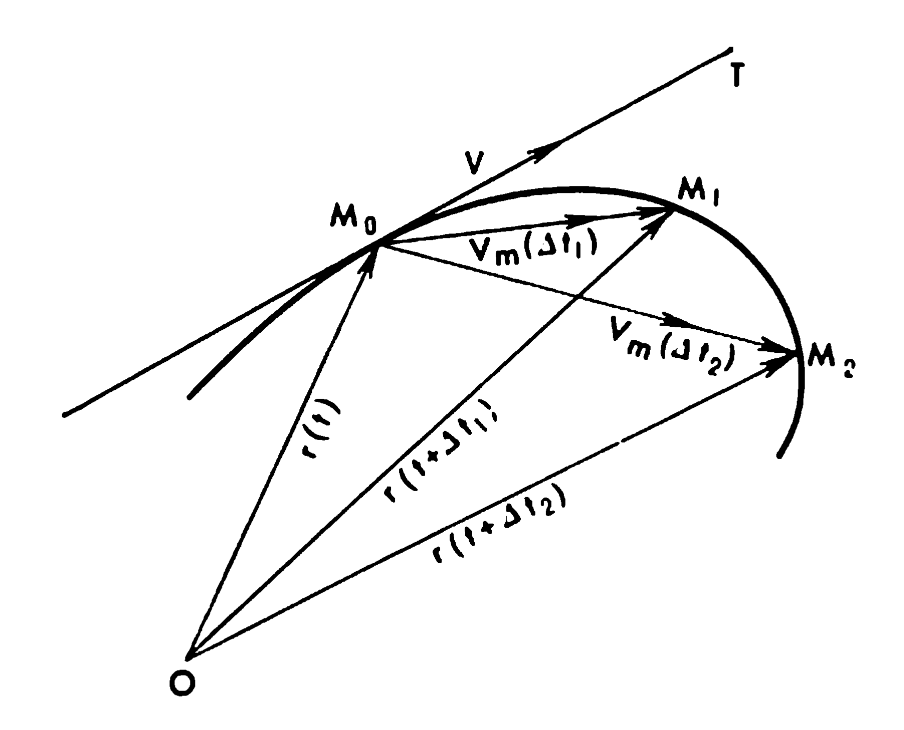


Fig. 30

Le vecteur v est appelé *vitesse instantanée du mouvement à l'instant* t_0. Son sens est le sens qu'auront à la limite les vecteurs vitesses moyennes.

Le vecteur vitesse moyenne pendant l'intervalle de temps $[t_0, t_1]$ est porté par la sécante M_0M_1. Si t_1 tend vers t_0, alors le point M_1 tendra vers M_0 en se déplaçant le long de la trajectoire. En tournant, la sécante M_0M_1 tendra vers une position limite M_0T appelée *tangente à la trajectoire au point* M_0 [1]). A l'instant t_0 le vecteur vitesse est porté par la tangente à la courbe au point M_0.

Le lecteur dont les connaissances ne sont pas très poussées peut être dérouté par les termes « tend », « limite », « position limite » qui, de surcroît, sont appliqués à des vecteurs variables (voire même à des droites variables) et pas seulement à des nombres. Les notions élémentaires de la théorie des limites, appliquée aux nombres variables, sont enseignées dans le secondaire. Nous aurons besoin,

---

[1]) On recommande au lecteur de comparer cette définition générale de la tangente avec la définition de la tangente au cercle.

nous, d'une théorie plus générale que nous allons exposer succinctement.

n°3. Soit $\rho(s)$ une fonction de s $(s > 0)$ à valeurs numériques [1]). Rappelons au lecteur la signification exacte de l'égalité

$$\lim_{s \to 0} \rho(s) = 0 \qquad (6)$$

que l'on lit : « la fonction $\rho(s)$ tend vers 0 lorsque s tend vers 0 ». On dit que la fonction $\rho(s)$ tend vers 0 lorsque s tend vers 0 si quel que soit un nombre $\varepsilon > 0$ arbitrairement petit, il existe un nombre $\delta > 0$ arbitrairement petit, tel que

$$|\rho(s)| < \varepsilon$$

pour tous les $s < \delta$.

Supposons maintenant que l'on ait une fonction vectorielle $\mathbf{a}(s)$. On dit que le vecteur $\mathbf{b}$ est limite de la fonction $\mathbf{a}(s)$ lorsque s tend vers 0, et l'on note $\mathbf{b} = \lim_{s \to 0} \mathbf{a}(s)$, si la fonction scalaire

$$\rho(s) = |\mathbf{a}(s) - \mathbf{b}|$$

tend vers 0 lorsque s tend vers 0.

Les principaux théorèmes de la théorie des limites sont les mêmes pour les fonctions vectorielles et les fonctions scalaires.

● THÉORÈME 1. *Si une fonction admet une limite, elle en admet une et une seule.*

● DÉMONSTRATION. Supposons qu'une fonction admette deux limites et prouvons qu'elles sont confondues. Soit

$$\mathbf{b}_1 = \lim_{s \to 0} \mathbf{a}(s) \quad \text{et} \quad \mathbf{b}_2 = \lim_{s \to 0} \mathbf{a}(s).$$

---

[1]) Ces fonctions sont appelées *scalaires* par opposition aux fonctions *vectorielles*.

Il est évident que

$$\mathbf{b}_1 - \mathbf{b}_2 = [\mathbf{b}_1 - \mathbf{a}(s)] + [\mathbf{a}(s) - \mathbf{b}_2].$$

D'où, en vertu de l'inégalité du triangle,

$$|\mathbf{b}_1 - \mathbf{b}_2| \leqslant |\mathbf{b}_1 - \mathbf{a}(s)| + |\mathbf{a}(s) - \mathbf{b}_2|.$$

Comme les deux termes du second membre tendent vers 0 lorsque $s \to 0$, et que le premier membre ne dépend pas de s, cette différence ne peut être positive:

$$|\mathbf{b}_1 - \mathbf{b}_2| \leqslant 0.$$

Or, elle ne peut être négative (la longueur d'un vecteur étant toujours positive). Donc

$$|\mathbf{b}_1 - \mathbf{b}_2| = 0.$$

D'où $\mathbf{b}_1 - \mathbf{b}_2 = 0$, ce qui entraîne $\mathbf{b}_1 = \mathbf{b}_2$.

● THÉORÈME 2. *Si*

$$\lim_{s \to 0} \mathbf{a}_1(s) = \mathbf{b}_1, \qquad \lim_{s \to 0} \mathbf{a}_2(s) = \mathbf{b}_2,$$

alors

$$\lim_{s \to 0} (\mathbf{a}_1(s) + \mathbf{a}_2(s)) = \mathbf{b}_1 + \mathbf{b}_2$$

« la limite d'une somme est égale à la somme des limites »).

● THÉORÈME 3. *Si*

$$\lim_{s \to 0} \mathbf{a}(s) = \mathbf{b},$$

alors quel que soit un nombre fixe λ *on a*

$$\lim_{s \to 0} \lambda \mathbf{a}(s) = \lambda \mathbf{b}$$

(« la limite du produit par un nombre est égal au produit de la limite par ce nombre »).

On laisse au lecteur le soin de démontrer les théorèmes 2 et 3 [1]).

[1]) On utilisera l'inégalité du triangle pour démontrer le théorème 2.

● THÉORÈME. 4. *Si*

$$\lim_{s\to 0} \mathbf{a}(s) = \mathbf{b},$$

alors quel que soit un angle fixe α on a

$$\lim_{s\to 0} U_\alpha \mathbf{a}(s) = U_\alpha \mathbf{b}.$$

● DÉMONSTRATION. On a (voir § 1, formule (20)):

$$| U_\alpha \mathbf{a}(s) - U_\alpha \mathbf{b} | = | U_\alpha (\mathbf{a}(s) - \mathbf{b}) | = | \mathbf{a}(s) - \mathbf{b} |.$$

Or, par hypothèse,

$$\lim_{s\to 0} | \mathbf{a}(s) - \mathbf{b} | = 0,$$

d'où

$$\lim_{s\to 0} | U_\alpha \mathbf{a}(s) - U_\alpha \mathbf{b} | = 0,$$

i.e.

$$\lim_{s\to 0} U_\alpha \mathbf{a}(s) = U_\alpha \mathbf{b}.$$

n°4. Désormais nous sommes en mesure d'exposer la théorie des vitesses dans la forme qui nous convient.

La vitesse d'un point est définie par l'égalité (5)[1]).

Il s'ensuit immédiatement

● THÉORÈME. 5. *La vitesse d'un point fixe est nulle* [2]).

● DÉMONSTRATION. En effet, si un point est fixe, quel que soit l'intervalle de temps considéré, le vecteur déplacement est nul. i. e. $\Delta \mathbf{r} = 0$.

Donc $\mathbf{v}_{moy} = \frac{\Delta \mathbf{r}}{\Delta t} = 0$ et $\mathbf{v} = \lim_{\Delta t\to 0} \mathbf{v}_{moy} = 0$ à chaque instant.

---

[1]) Le lecteur familiarisé aux notions de dérivée dirait: « la vitesse d'un point est égale à la dérivée de son rayon vecteur par rapport au temps ». La dérivation est une opération qui joue un très grand rôle en mathématiques. Au lecteur intéressé on recommande la brochure de V. Boltianski « Qu'est ce que la dérivation? » (Gostekhizdat, Moscou, 1955).

[2]) Le temps bien sûr que ce point reste fixe.

Formulons la réciproque.

● THÉORÈME. 5'. *Si la vitesse d'un point est nulle (le temps qu'est étudié le mouvement de ce point), alors il est fixe.*

Quoique physiquement évident, ce théorème n'est pas si simple à démontrer mathématiquement. Pour ne pas trop nous écarter du sujet nous l'admettrons sans le démontrer.

Les théorèmes 5 et 5' indiquent que l'égalité

$$\mathbf{r} = \text{const}$$

est équivalente à $\mathbf{v} = 0$.

● THÉORÈME. 6. *Soient* $\mathbf{r}_1 = \mathbf{r}_1(t)$, $\mathbf{r}_2 = \mathbf{r}_2(t)$ *et* $\mathbf{r} = \mathbf{r}(t)$ *les rayons vecteurs des points* M_1, M_2 *et* M *respectivement. Si ces points son animés de mouvements tels que l'on ait toujours la relation*

$$\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2,$$

alors leurs vitesses sont liées par la relation analogue:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2. \qquad (7)$$

● DÉMONSTRATION. Le vecteur déplacement du point M pendant l'intervalle de temps $[t, t + \Delta t]$ est égal à

$$\Delta\mathbf{r} = \mathbf{r}(t + \Delta t) - \mathbf{r}(t) = [\mathbf{r}_1(t + \Delta t) + \mathbf{r}_2(t + \Delta t)] - [\mathbf{r}_1(t) + \mathbf{r}_2(t)] = [\mathbf{r}_1(t + \Delta t) - \mathbf{r}_1(t)] + [\mathbf{r}_2(t + \Delta t) - \mathbf{r}_2(t)] = \Delta\mathbf{r}_1 + \Delta\mathbf{r}_2.$$

D'où

$$\mathbf{v}_{\text{moy}} = \frac{\Delta\mathbf{r}}{\Delta t} = \frac{\Delta\mathbf{r}_1}{\Delta t} + \frac{\Delta\mathbf{r}_2}{\Delta t} = \mathbf{v}_{1\,\text{moy}} + \mathbf{v}_{2\,\text{moy}}$$

et, d'après le théorème 2,

$$\mathbf{v} = \lim_{\Delta t \to 0} \mathbf{v}_{\text{moy}} = \lim_{\Delta t \to 0} \mathbf{v}_{1\,\text{moy}} + \lim_{\Delta t \to 0} \mathbf{v}_{2\,\text{moy}} = \mathbf{v}_1 + \mathbf{v}_2,$$

ce qu'il fallait démontrer.

On démontrerait le même théorème pour la différence.

● THÉORÈME 6′. *Si les vitesses des points* M_1, M_2 *et* M *vérifient la relation* (7), *alors*

$$\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2 + \text{const.} \qquad (8)$$

● DÉMONSTRATION. Considérons le point auxiliaire P dont le rayon vecteur est toujours égal à

$$\overrightarrow{OP} = \mathbf{r} - (\mathbf{r}_1 + \mathbf{r}_2). \qquad (9)$$

De ce qui précède, la vitesse du point P est donc égale à $\mathbf{v} - (\mathbf{v}_1 + \mathbf{v}_2)$, i. e. elle est nulle. Le point P est donc fixe, d'où

$$\overline{OP} = \text{const.}$$

Cette relation et la relation (9) entraînent (8).

En utilisant les théorèmes 3 et 4, on démontrerait les théorèmes suivants d'une façon analogue aux théorèmes 6 et 6′:

● THÉORÈME 7. *Soient* $\mathbf{r}_1 = \mathbf{r}_1(t)$ *et* $\mathbf{r}_2 = \mathbf{r}_2(t)$ *les rayons vecteurs respectivement des points* M_1 *et* M_2. *Si ces points sont animés de mouvements tels que l'on ait toujours*

$$\mathbf{r}_2 = \lambda \mathbf{r}_1,$$

où λ *est un nombre constant, alors leurs vitesses sont liées par la relation analogue*

$$\mathbf{v}_2 = \lambda \mathbf{v}_1. \qquad (10)$$

● THÉORÈME 7′. *Si les vitesses des points* M_1 *et* M_2 *vérifient la relation* (10), *alors*

$$\mathbf{r}_2 = \lambda \mathbf{r}_1 + \text{const.}$$

● THÉORÈME 8. *Soient* $\mathbf{r}_1 = \mathbf{r}_1(t)$ *et* $\mathbf{r}_2 = \mathbf{r}_2(t)$ *les rayons vecteurs respectivement des points* M_1 *et* M_2. *Si ces points sont animés de mouvements tels que*

$$\mathbf{r}_2 = U_\alpha \mathbf{r}_1,$$

où α est un angle constant, alors leurs vitesses sont liées par la relation analogue

$$\mathbf{v}_2 = U_\alpha \mathbf{v}_1. \tag{11}$$

● THÉORÈME 8'. *Si les vitesses des points M_1 et M_2 sont liées par la relation* (11), *alors*

$$\mathbf{r}_2 = U_\alpha \mathbf{r}_1 + \text{const}.$$

n°5. Soit $\mathbf{r} = \mathbf{r}(t)$ le rayon vecteur du point mobile M. Considérons le déplacement $\overrightarrow{M_0M_1} = \Delta\mathbf{r}$ de ce point pendant l'intervalle de temps $[t_0, t_1]$ (fig. 31). Décrivons un arc de cercle de centre O et de rayon OM_0. Il coupe le rayon $\overrightarrow{OM_1}$ en M^*. Il est évident que

$$\overrightarrow{\Delta\mathbf{r}} = \overrightarrow{M_0M_1} = \overrightarrow{M_0M^*} + \overrightarrow{M^*M_1}. \tag{12}$$

Le déplacement $\overrightarrow{M_0M^*}$ laisse invariante la distance du point mobile M au pôle O, il n'est lié qu'à la rotation du rayon $\overrightarrow{OM}$. Le déplacement $\overrightarrow{M^*M_1}$ est lié uniquement aux variations de la distance du point M au pôle O.

En divisant les deux membres de l'égalité (12) par $\Delta t = t_1 - t_0$ et en passant à la limite lorsque $\Delta t \to 0$, on obtient:

$$\mathbf{v} = \lim_{\Delta t \to 0} \frac{\overrightarrow{M_0M^*}}{\Delta t} + \lim_{\Delta t \to 0} \frac{\overrightarrow{M^*M_1}}{\Delta t}. \tag{13}$$

La première limite est appelée *vitesse transversale* du point M et notée $\mathbf{v}_t$, la deuxième vitesse *radiale* du point M est notée $\mathbf{v}_r$. On a donc

$$\mathbf{v} = \mathbf{v}_t + \mathbf{v}_r. \tag{14}$$

La formule (14) permet de décomposer le vecteur vitesse en une vitesse radiale et une vitesse transversale (fig. 32). Ces composantes sont perpendiculaires.

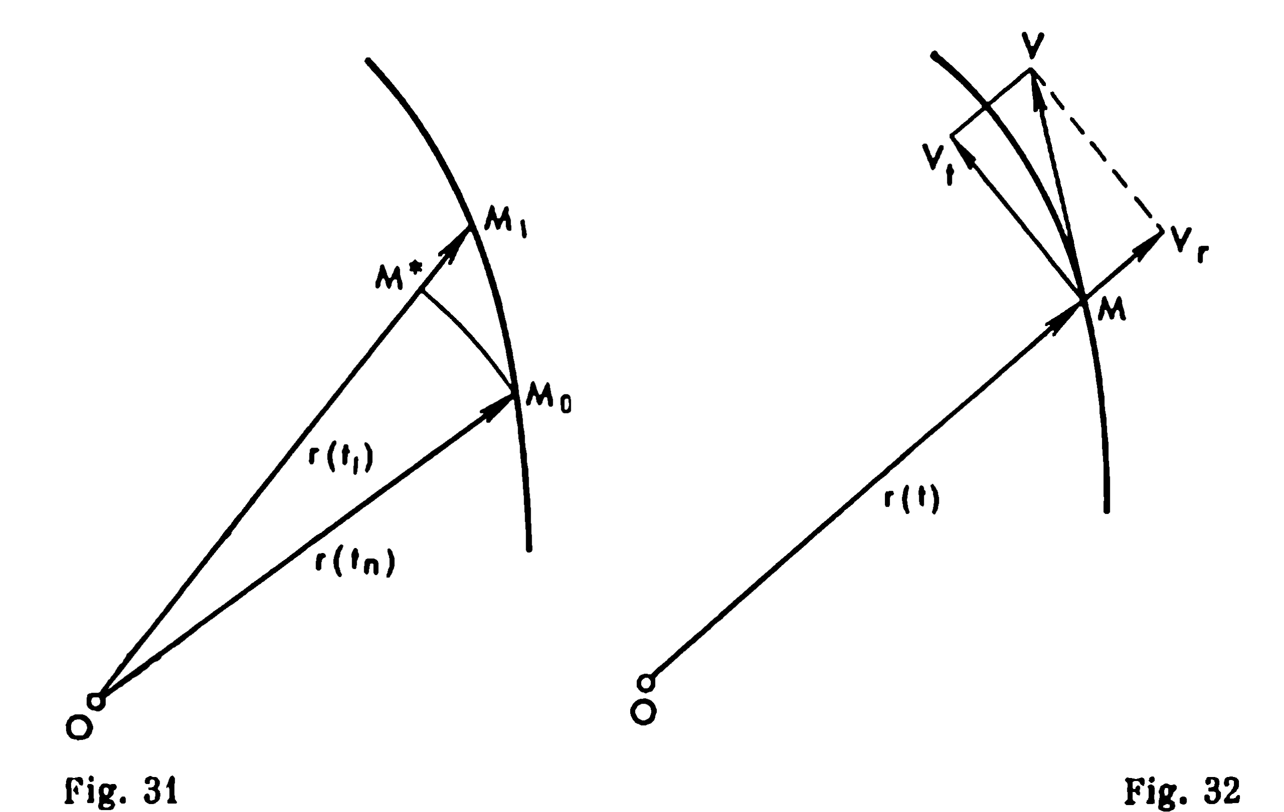


Fig. 31 Fig. 32

La vitesse radiale est la vitesse de variation de la distance du point M au pôle O ou, ce qui revient au même, la vitesse de variation du module du rayon vecteur $\overrightarrow{OM}$. Elle est orientée dans le sens de ce vecteur si $\overrightarrow{OM}$ croît et dans le sens contraire si $\overrightarrow{OM}$ décroît.

Soit $\mathbf{v}_r$ la projection du vecteur vitesse du point M sur le support du vecteur $\overrightarrow{OM}$. Il est évident que

$$\mathbf{v}_r = \pm \mid \mathbf{v}_r \mid,$$

où l'on prend le signe plus lorsque $\overrightarrow{OM}$ croît et le signe moins lorsque $\overrightarrow{OM}$ décroît.

Si le point M décrit un cercle centré en O, sa vitesse totale coïncide avec sa vitesse transversale, i. e.

$$\mathbf{v} = \mathbf{v}_t, \qquad \mathbf{v}_r = 0.$$

S'il décrit un rayon issu du pôle O, sa vitesse totale coïncide avec sa vitesse radiale

$$\mathbf{v} = \mathbf{v}_r, \qquad \mathbf{v}_t = 0.$$

n°6. Considérons la rotation du rayon $\overrightarrow{OM}$ autour de son origine O. Supposons que durant l'intervalle $[t, t + \Delta t]$ le rayon ait pivoté d'un angle [1] $\Delta\varphi$. On appelle *vitesse angulaire moyenne* du rayon $\overrightarrow{OM}$ pendant l'intervalle de temps $[t, t + \Delta t]$ le rapport

$$\omega_{moy} = \frac{\Delta\varphi}{\Delta t}.$$

On appelle *vitesse angulaire* (*instantanée*) du rayon $\overrightarrow{OM}$ et on la désigne par ω la limite de la vitesse angulaire moyenne ω_{moy} lorsque $\Delta t \to 0$:

$$\omega = \lim_{\Delta t \to 0} \omega_{moy} = \lim_{\Delta t \to 0} \frac{\Delta\varphi}{\Delta t}.$$

La vitesse angulaire n'est pas un vecteur mais un nombre [2]. Elle a le signe plus ou moins selon que la rotation du rayon s'opère dans le sens positif ou négatif.

Pour les vitesses angulaires d'un rayon ont lieu les théorèmes suivants qui rappellent les théorèmes sur les vitesses d'un point.

● THÉORÈME 9. *La vitesse angulaire d'un rayon est constamment nulle si, et seulement si, ce rayon est constamment fixe.*

● THÉORÈME 10. *Les vitesses angulaires de deux rayons* [3] *O_1M et O_2M sont constamment égales, si, et seulement si l'angle qu'ils font est constant.*

---

[1] Cet angle peut être positif ou négatif.

[2] Pour les mouvements non plans la notion de vitesse angulaire est plus complexe, et souvent est une grandeur vectorielle.

[3] O_1 et O_2 sont deux points arbitraires fixes (qui peuvent bien être confondus).

§ 3 MÉTHODE CINÉMATIQUE DANS LES PROBLÈMES DE GÉOMÉTRIE

Nous voilà suffisamment armés pour aborder les problèmes de géométrie. Auparavant, nous recommanderions une seconde analyse de la résolution cinétique du « problème du trésor » examiné dans l'introduction, et une étude approfondie des notions exposées dans les §§ 1 et 2.

● PROBLÈME 1[1]). *Sur les côtés et à l'extérieur d'un triangle quelconque ABC on construit trois triangles équilatéraux ABC', BCA' et ACB'* (fig. 33). *On demande de prouver que les centres O_1, O_2 et O_3 de ces triangles sont les sommets d'un triangle équilatéral.*

● SOLUTION. Fixons les sommets A et B du triangle ABC et faisons déplacer le sommet C à une vitesse v_c. Le triangle ABC' sera fixe, tandis que les sommets A' et B' des triangles équilatéraux $A'BC$ et $AB'C$ seront animés d'un certain mouvement. Soient les vecteurs $\overrightarrow{AC}$ et $\overrightarrow{AO_2}$. Il est évident que

$$AO_2 = \frac{1}{\sqrt{3}} AC.$$

D'autre part l'angle des vecteurs $\overrightarrow{AC}$ et $\overrightarrow{AO_2}$ est égal à $\frac{\pi}{6}$. Si donc nous faisons pivoter le vecteur $\overrightarrow{AC}$ d'un angle de $\frac{\pi}{6}$ (sans modifier sa longueur) et que nous multiplions le vecteur ainsi obtenu par $\frac{1}{\sqrt{3}}$, nous déduisons le vecteur

---

[1]) Ce problème, autant d'ailleurs que la plupart des problèmes proposés dans la suite, est emprunté à la série « Bibliothèque du cercle de mathématiques » de I. Yaglom et autres. Certains problèmes sont empruntés à la « Géométrie élémentaire » de G. Adamard, t. 1.

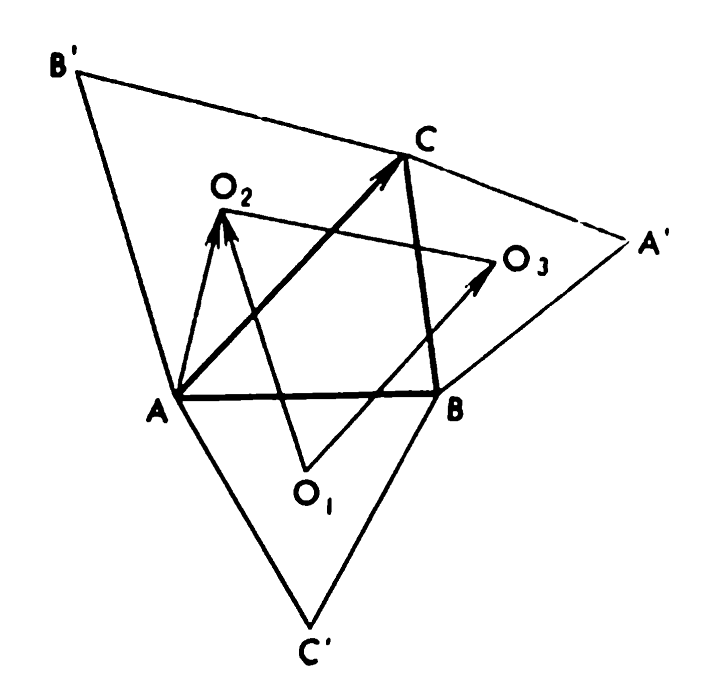


Fig. 33

$\overrightarrow{AO_2}$. Ce que l'on note

$$\overrightarrow{AO_2} = \frac{1}{\sqrt{3}} U_{\frac{\pi}{6}} \overrightarrow{AC}.$$

Or, d'après le théorème 8,

$$\mathbf{v}_{O_2} = \frac{1}{\sqrt{3}} U_{\frac{\pi}{6}} \mathbf{v}_c$$

($\mathbf{v}_{O_2}$ est la vitesse du point O_2).
D'une façon analogue

$$\mathbf{v}_{O_2} = \frac{1}{\sqrt{3}} U_{-\frac{\pi}{6}} \mathbf{v}_C.$$

Donc

$$\mathbf{v}_C = \sqrt{3}\, U_{\frac{\pi}{6}} \mathbf{v}_{O_2}.$$

Et

$$\mathbf{v}_{O_2} = \frac{1}{\sqrt{3}} U_{\frac{\pi}{6}} \sqrt{3}\, U_{\frac{\pi}{6}} \mathbf{v}_{O_2} = U_{\frac{\pi}{6}} U_{\frac{\pi}{6}} \mathbf{v}_{O_2} = U_{\frac{\pi}{3}} \mathbf{v}_{O_2}. \qquad (1)$$

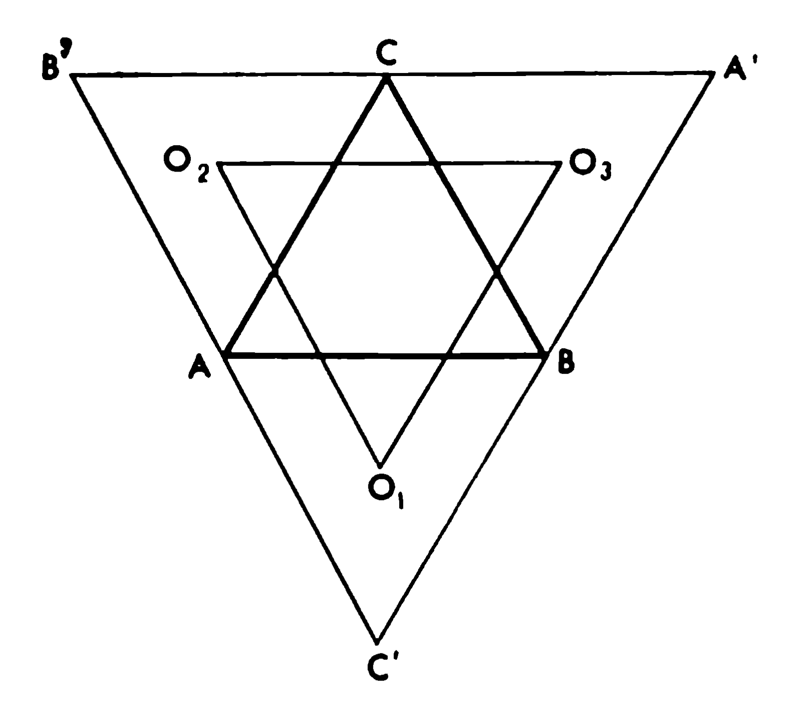


Fig. 34

Prenons maintenant le point fixe O_1 pour pôle. En appliquant le théorème 8′ à l'égalité (1) il vient alors:

$$\overrightarrow{O_1O_2} = U_{\frac{\pi}{3}}\overrightarrow{O_1O_3} + \mathbf{R},$$

où $\mathbf{R}$ = const, i. e. le vecteur $\mathbf{R}$ ne dépend pas de la position du point mobile C. Le vecteur $\mathbf{R}$ est inconnu, mais on peut le trouver en fixant une position du point C que nous appellerons *position déterminante*. S'il s'avère que dans une position déterminante du point C, le vecteur $\mathbf{R}$ est nul, cela veut dire, puisqu'il est constant, qu'il est toujours nul, i. e.

$$\overrightarrow{O_1O_2} = U_{\frac{\pi}{3}}\overrightarrow{O_1O_3}. \qquad (2)$$

Or, cette relation signifie précisément que le triangle $O_1O_2O_3$ est constamment équilatéral!

En effet, la relation (2) indique que le segment O_1O_2 se déduit du segment O_1O_3 par une rotation de $\frac{\pi}{3}$.

Il reste à trouver une position déterminante convenable du point C. On s'arrangera pour que la figure soit la plus simple possible. Dans notre problème, on obtient ce résultat si le point C occupe une position telle (fig. 34) que le triangle ABC soit équilatéral. Dans ce cas, la figure est symétrique puisqu'elle vient en coïncidence avec elle-même si elle pivote de $\frac{2}{3}\pi$ autour du centre du triangle ABC. Le triangle $O_1O_2O_3$ est donc équilatéral et

$$\overrightarrow{O_1O_2} = U_{\frac{\pi}{3}}\overrightarrow{O_1O_3},$$

i. e. dans cette position $\mathbf{R} = 0$.

● EXERCICES.

1. Montrer que la proposition du problème 1 reste valable si on remplace les triangles ABC', BCA' et ACB' par les triangles ABC'', BCA'' et ACB'' qui leur sont symétriques par rapport aux côtés du triangle ABC (fig. 35).

2. Sur les côtés d'un triangle quelconque ABC, on construit les triangles équilatéraux BCA', ACB' et ABC' de telle sorte que les sommets A' et A, B et B' soient situés de part et d'autre respectivement des côtés BC et AC, et les sommets C' et C du même côté de AB (fig. 36). Montrer que si le point M est le centre du triangle ABC', le triangle $A'MB'$ est isocèle et son angle au sommet M est égal à $\frac{2}{3}\pi$.

3. Sur les côtés et à l'extérieur d'un triangle quelconque ABC on construit les triangles isocèles BCA', ACB' et ABC' d'angles au sommet respectivement α, β et γ (fig. 37). Montrer que si

$$\alpha + \beta + \gamma = 2\pi,$$

les angles du triangle $A'B'C'$ sont égaux à $\frac{\alpha}{2}$, $\frac{\beta}{2}$, $\frac{\gamma}{2}$, i.e. ne dépendent pas de la forme du triangle ABC.

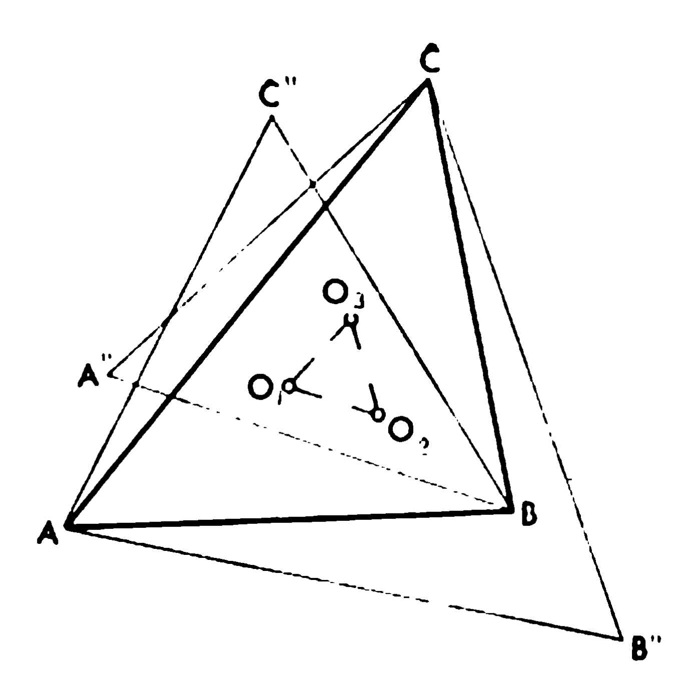


Fig. 35

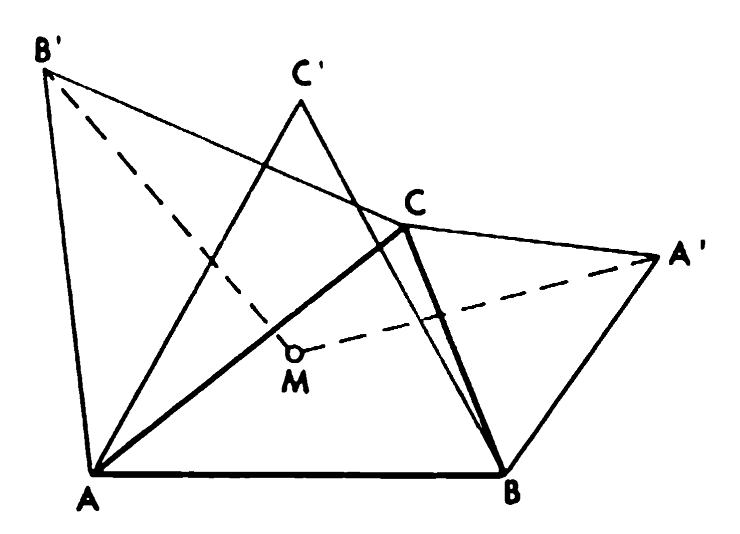


Fig. 36

● PROBLÈME 2. *Sur les côtés et à l'extérieur d'un quadrilatère $ABCD$, on construit les triangles rectangles isocèles ABM, BCP, CDQ et DAS* (fig. 38). *On demande de prouver que les segments MQ et SP sont égaux et orthogonaux.*

● SOLUTION. Fixons les sommets A, B et D et déplaçons le sommet C. Comme

$$\overrightarrow{BP} = \frac{1}{\sqrt{2}} U_{\frac{\pi}{4}} \overrightarrow{BC},$$

alors

$$\mathbf{v}_P = \frac{1}{\sqrt{2}} U_{\frac{\pi}{4}} \mathbf{v}_C.$$

D'une façon analogue

$$\mathbf{v}_Q = \frac{1}{\sqrt{2}} U_{-\frac{\pi}{4}} \mathbf{v}_C.$$

En passant aux rayons vecteurs on obtient

$$\overrightarrow{SP} = U_{\frac{\pi}{2}} \overrightarrow{SQ} + \text{const}.$$

Comme

$$\overrightarrow{SQ} = \overrightarrow{SM} + \overrightarrow{MQ}$$

et

$$\overrightarrow{SM} = \text{const},$$

il vient

$$\overrightarrow{SP} = U_{\frac{\pi}{2}} \overrightarrow{MQ} + \mathbf{R},$$

où

$$\mathbf{R} = \text{const}.$$

Faisons choix d'une position déterminante du sommet C. Faisons-le coïncider par exemple avec le sommet A. Le quadrilatère $ABCD$ se transforme en deux paires de segments confondus: $AB = CB$ et $AD = CD$ (fig. 39). Les triangles ABM et CBP forment un carré de diagonale AB. Il en est de même des triangles ADS et CDQ qui déterminent un carré de diagonale AD. D'où il résulte qu'une rotation d'angle $\frac{\pi}{2}$ amènera le triangle ASP en coïncidence avec le triangle AQM (le sommet S venant en Q, le sommet P

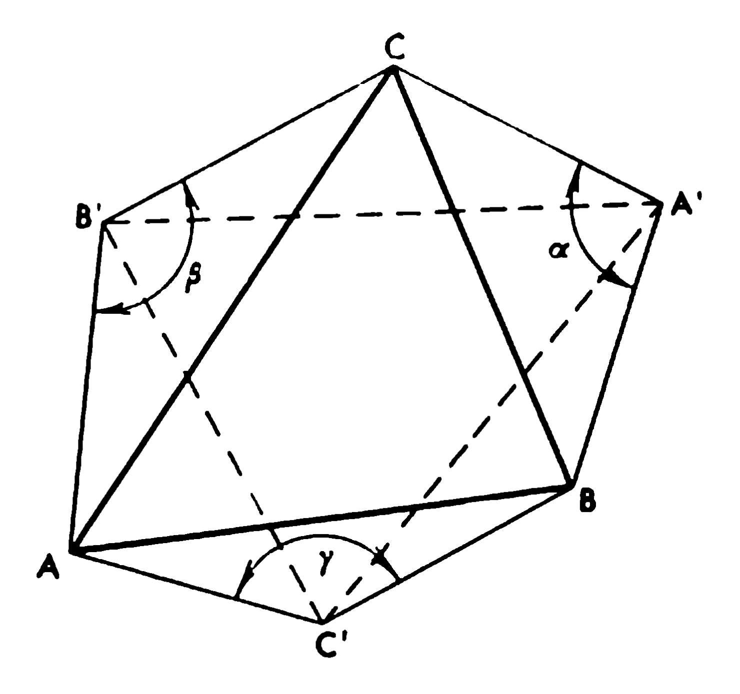


Fig. 37

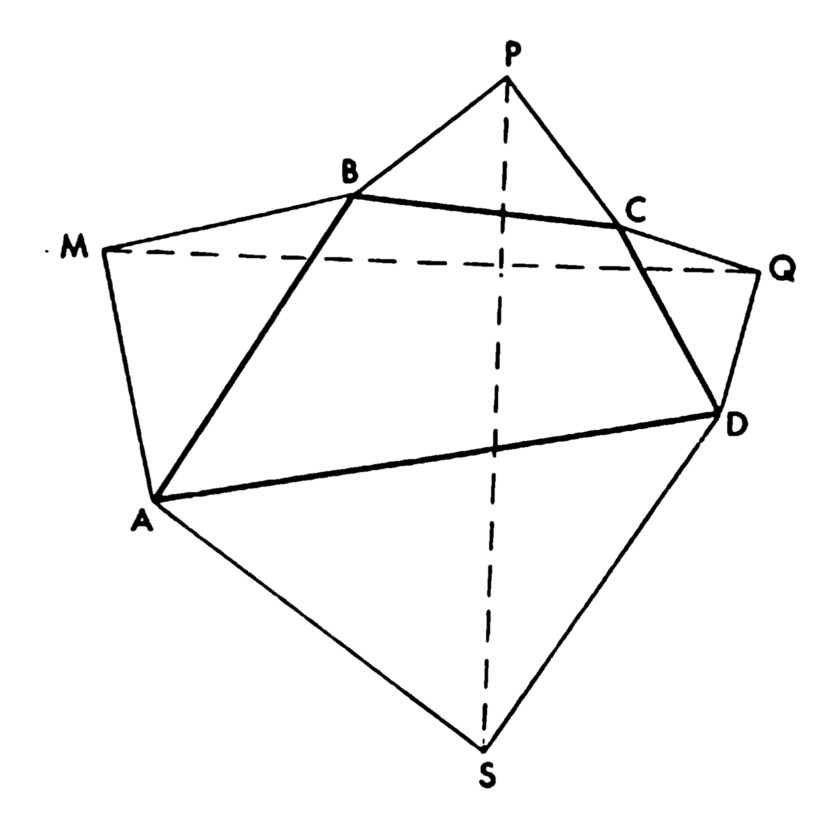


Fig. 38

en M). Dans cette position déterminante du point C, nous avons donc

$$\overrightarrow{SP}=U_{\frac{\pi}{2}}\overrightarrow{MQ}.$$

D'où l'on déduit que

$$\mathbf{R}=0,$$

et on a toujours

$$\overrightarrow{SP}=U_{\frac{\pi}{2}}\overrightarrow{MQ}.$$

Or, cette relation exprime le fait que le segment SP se déduit du segment MQ par une rotation d'angle $\frac{\pi}{2}$. Donc

$$SP = MQ, \qquad SP \perp MQ.$$

● PROBLÈME 3. *Sur les côtés et à l'extérieur d'un parallélogramme $ABCD$, on construit quatre carrés. On demande de prouver que les centres M, P, Q, S de ces carrés sont les sommets d'un carré* (fig. 40).

● SOLUTION. Fixons les points A et D et animons le segment BC d'un mouvement de translation. Remarquons que les points B et C se déplaceront à la même vitesse, de même d'ailleurs que le centre Q du carré construit sur le côté BC.

Calculons la vitesse $\mathbf{v}_S$ du point S, centre du carré construit sur le côté CD. Comme

$$\overrightarrow{DS}=\frac{1}{\sqrt{2}}U_{\frac{\pi}{4}}\overrightarrow{DC},$$

il vient

$$\mathbf{v}_S=\frac{1}{\sqrt{2}}U_{\frac{\pi}{4}}\mathbf{v}_C.$$

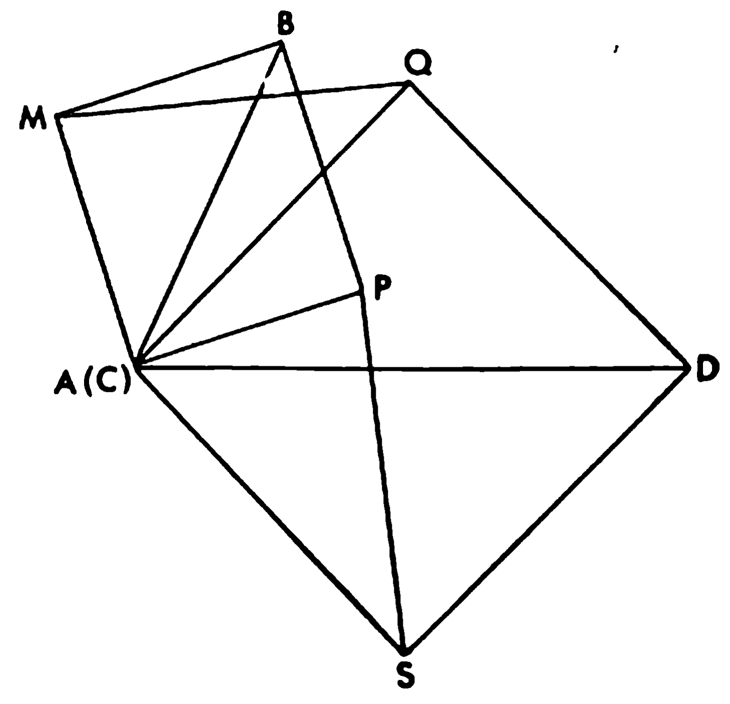


Fig. 39

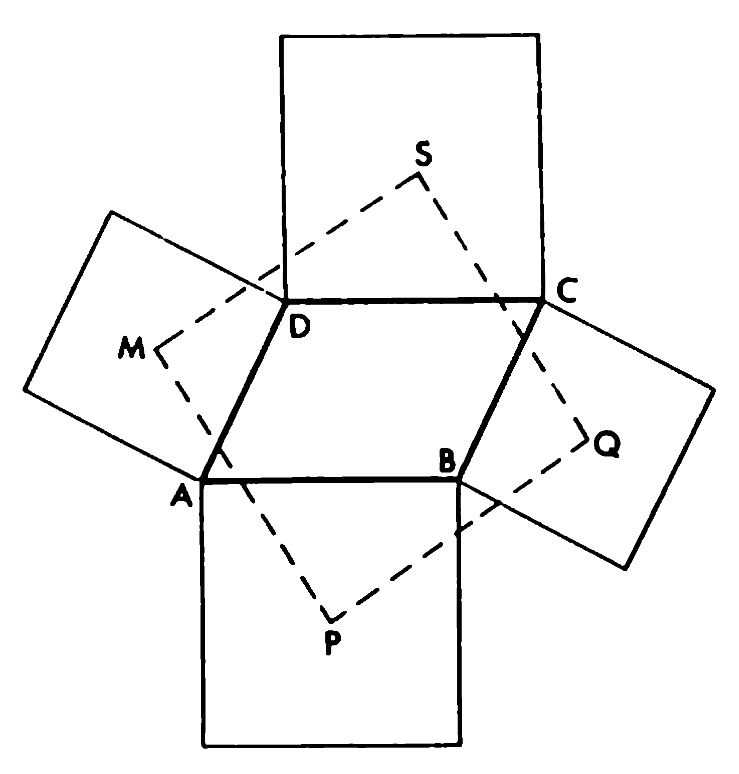


Fig. 40

D'une façon analogue:

$$\mathbf{v}_P = \frac{1}{\sqrt{2}} U_{-\frac{\pi}{4}} \mathbf{v}_B.$$

Puisque

$$\mathbf{v}_B = \mathbf{v}_C = \mathbf{v}_Q,$$

il vient

$$\mathbf{v}_S = \frac{1}{\sqrt{2}} U_{\frac{\pi}{4}} \mathbf{v}_Q, \quad \mathbf{v}_P = \frac{1}{\sqrt{2}} U_{-\frac{\pi}{4}} \mathbf{v}_Q$$

et

$$\overrightarrow{MS} = \frac{1}{\sqrt{2}} U_{\frac{\pi}{4}} \overrightarrow{MQ} + \mathbf{R}_1, \quad \overrightarrow{MP} = \frac{1}{\sqrt{2}} U_{-\frac{\pi}{4}} \overrightarrow{MQ} + \mathbf{R}_2,$$

où

$$\mathbf{R}_1 = \text{const}, \quad \mathbf{R}_2 = \text{const}.$$

Pour position déterminante, nous allons prendre une position du segment BC, telle que le parallélogramme $ABCD$ soit un carré. On établit alors que

$$\mathbf{R}_1 = 0, \qquad \mathbf{R}_2 = 0.$$

Donc, on a toujours

$$\overrightarrow{MS} = \frac{1}{\sqrt{2}} U_{\frac{\pi}{4}} \overrightarrow{MQ}, \quad \overrightarrow{MP} = \frac{1}{\sqrt{2}} U_{-\frac{\pi}{4}} \overrightarrow{MQ},$$

ce qui prouve que $MPQS$ est un carré.

● EXERCICES.

4. Montrer que la proposition du problème 3 reste valable si l'on y remplace tous les carrés par des carrés symétriques par rapport aux côtés du parallélogramme $ABCD$.

5. On donne un quadrilatère $ABCD$. Montrer que si les sommets P et S des triangles rectangles isocèles ABP et CDS sont confondus, alors les sommets Q et T des triangles rectangles isocèles BCQ et DAT sont également confondus (fig. 41). Tous les triangles sont construits à l'intérieur du quadrilatère $ABCD$.

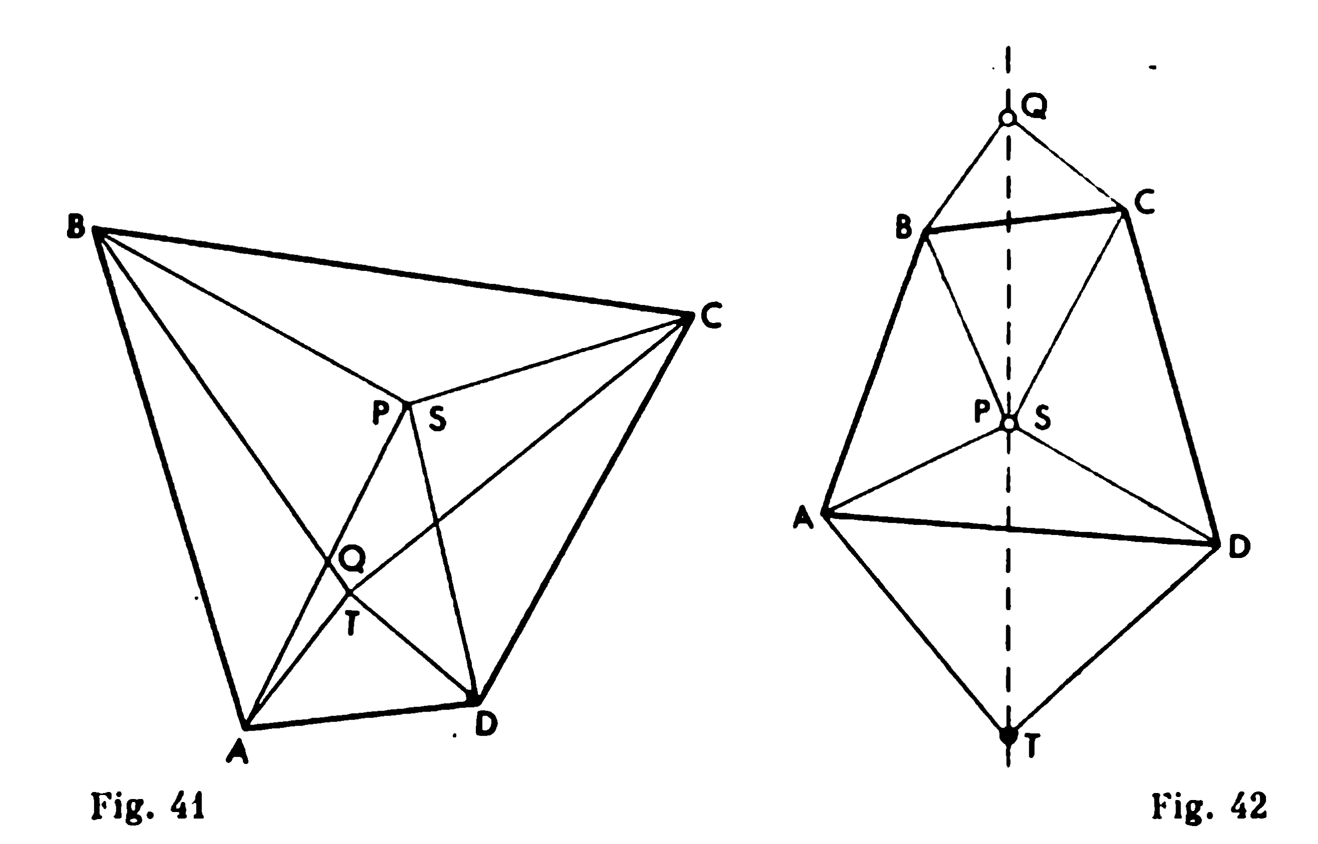


Fig. 41 Fig. 42

6. On donne un quadrilatère $ABCD$ et on construit les triangles rectangles isocèles BCQ et ADT à l'extérieur et les triangles rectangles isocèles CDS et ABP à l'intérieur (fig. 42). Montrer que si les sommets P et S sont confondus, le segment QT passe par ces sommets et les admet pour milieu.

7. Montrer que dans le problème 1, les segments AA', BB' et CC' sont égaux, concourants et forment des angles de $\frac{2}{3}\pi$.

● PROBLÈME 4. *On donne quatre droites* (a), (b), (c), (d) *qui se coupent deux à deux en six points* A, B, C, D, E, F (fig. 43). *Prouver que les milieux* M, P *et* Q *des segments* AC, BE *et* DF *sont alignés.*

● SOLUTION. Fixons les droites (b), (c) et (d) et animons la droite (a) d'un mouvement de translation. Les points B, C et D se déplaceront respectivement sur les droites (b), c) et (d) à des vitesses v_B, v_C et v_D.

Soient (a'), B', C' et D' les nouvelles positions de la droite (a) et des points B, C et D (fig. 44). Les origines et les

extrémités des vecteurs $\overrightarrow{BB'}$, $\overrightarrow{CC'}$ et $\overrightarrow{DD'}$ sont portées par les droites parallèles (a) et (a'). Si donc, on ramène les origines de ces vecteurs à un même point, leurs extrémités seront situées sur une droite parallèle à la droite (a). Les vecteurs $\overrightarrow{BB'}$, $\overrightarrow{CC'}$ et $\overrightarrow{DD'}$ sont proportionnels aux vitesses $\mathbf{v}_B$, $\mathbf{v}_C$ et $\mathbf{v}_D$ des points B, C et D. Si donc, comme on vient de le dire, on ramène l'origine de ces trois vecteurs à un même point, leurs extrémités B_1, C_1 et D_1 seront situées sur une droite parallèle à la droite (a). D'où il résulte, en vertu du théorème du § 1, n° 7, qu'il existe des constantes m et n telles que

$$\mathbf{v}_B = \frac{m\mathbf{v}_C + n\mathbf{v}_D}{m+n}. \tag{3}$$

Le point M(voir fig. 43) étant situé au milieu du segment AC

$$\overrightarrow{AM} = \frac{1}{2}\overrightarrow{AC}.$$

Comme le point A est fixe, il vient

$$\mathbf{v}_M = \frac{1}{2}\mathbf{v}_C. \tag{4}$$

D'une façon analogue, on obtient

$$\mathbf{v}_Q = \frac{1}{2}\mathbf{v}_D, \quad \mathbf{v}_P = \frac{1}{2}\mathbf{v}_B. \tag{5}$$

En vertu de (3)-(5), on a

$$\mathbf{v}_P = \frac{m\mathbf{v}_M + n\mathbf{v}_Q}{m+n}.$$

Prenons maintenant le point fixe E pour pôle. En utilisant successivement les théorèmes 6′ et 7′, on obtient

$$\overrightarrow{EP} = \frac{m\overrightarrow{EM} + n\overrightarrow{EQ}}{m+n} + \mathbf{R}, \tag{6}$$

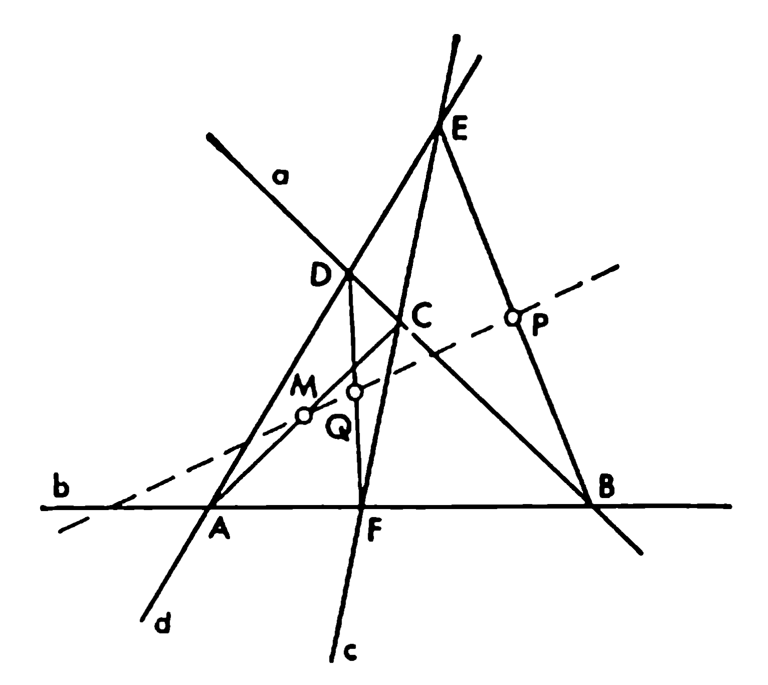


Fig. 43

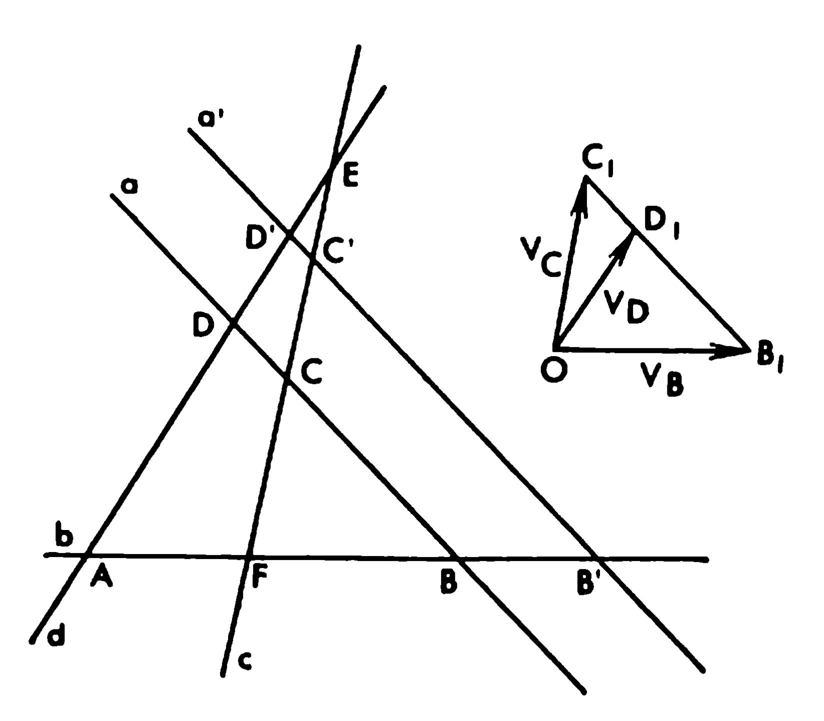


Fig. 44

où

$$\mathbf{R} = \text{const}.$$

Prenons une position déterminante de la droite (a) et supposons qu'elle passe par le point E (fig. 45). Dans cette position, les points D et C sont confondus avec le point E. Les points M, Q et P sont situés aux milieux des segments AE, FE et BE. Comme les points A, F et B sont portés par une même droite, les points M, Q et P seront également portés par une même droite (parallèle à (b)). Les triangles PEQ et B_1C_1O sont semblables, car leurs côtés sont respectivement parallèles ($PE \parallel a \parallel B_1C_1$; $PQ \parallel b \parallel OB_1$; $EQ \parallel c \parallel OC_1$).

D'une façon analogue, les triangles PEM et B_1D_1O sont semblables. Ceci entraîne

$$\frac{B_1D_1}{OB_1} = \frac{PE}{MP}, \qquad \frac{C_1B_1}{OB_1} = \frac{PE}{PQ},$$

d'où

$$\frac{MP}{PQ} = \frac{C_1B_1}{B_1D_1},$$

i.e. les points P et B_1 divisent respectivement les segments MQ et C_1D_1 dans le même rapport. Or, pour le point B ce rapport est $m : n$. Donc pour le point P, il est égal à $m : n$ et

$$\overrightarrow{EP} = \frac{m\overrightarrow{EM} + n\overrightarrow{EQ}}{m+n}. \tag{7}$$

En comparant les égalités (6) et (7), nous constatons que dans la position déterminante, $\mathbf{R} = 0$. Or, comme $\mathbf{R} = \text{const}$, il est donc toujours nul. Ce qui signifie qu'a lieu la relation (7) et les points M, P et Q sont toujours alignés.

● EXERCICE 8. Montrer que les orthocentres des triangles BCD, ABE, DEF, ACF obtenus par l'intersection deux à deux des droites (a), (b), (c) et (d) sont alignés (fig. 46).

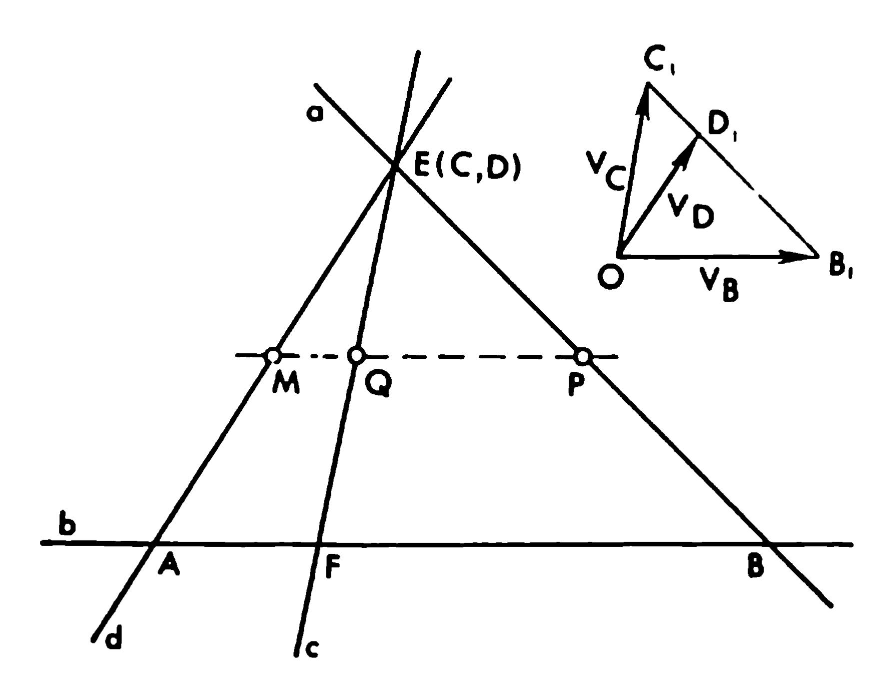


Fig. 45

● PROBLÈME 5. *Soient P un point quelconque du cercle (K) circonscrit au triangle ABC, et P_1, P_2, P_3 ses projections respectivement sur les côtés AB, AC et BC* (fig. 47). *Montrer que les points P_1, P_2, P_3 sont alignés (la droite qui passe par ces trois points est appelée droite de Simpson du point P du cercle circonscrit au triangle ABC).*

● SOLUTION. Faisons pivoter les côtés AC et BC autour des points A et B à une même vitesse angulaire ω. Le point C décrira le cercle (K)[1]. Comme les angles PP_1A et PP_2A sont droits, le point P_2 décrira le cercle K_1 passant par les points fixes A, P et P_1. De plus, le rayon PP_2, qui reste constamment perpendiculaire au rayon AC, tourne autour du point P à la même vitesse angulaire ω (en vertu du théorème 10 § 2). Puisque les rayons PP_2 et P_1P_2 sont animés d'un mouvement de rotation, tel que leur point de concours P_2 décrit le cercle (K_1), l'angle qu'ils forment est constant.

[1]) En effet, les vitesses angulaires des rayons AC et BC étant égales, en vertu du théorème 10, l'angle ACB est constant et le point C décrit donc le cercle (K).

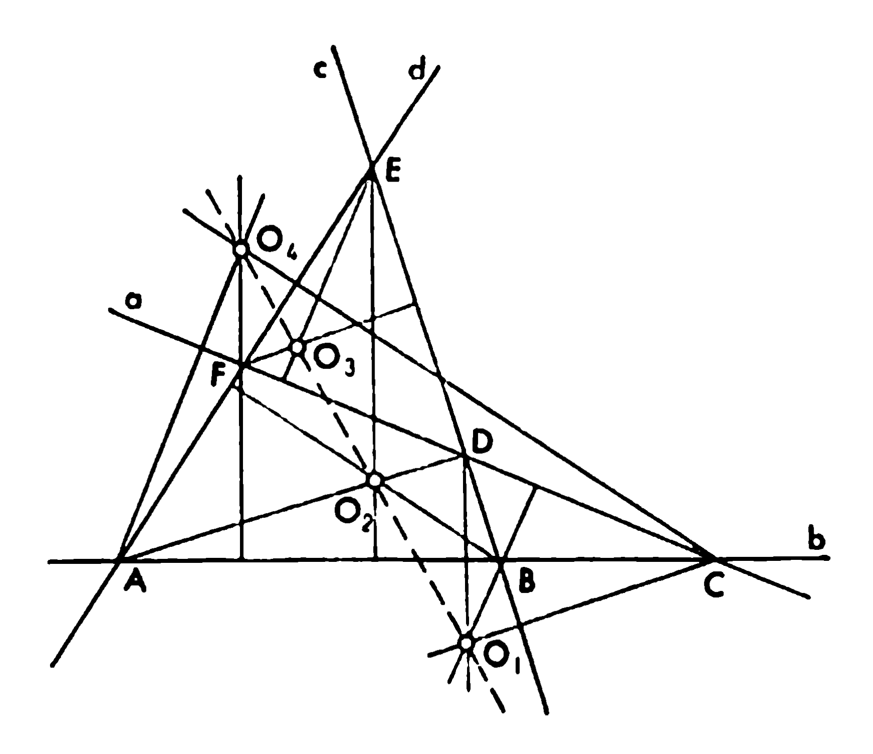


Fig. 46

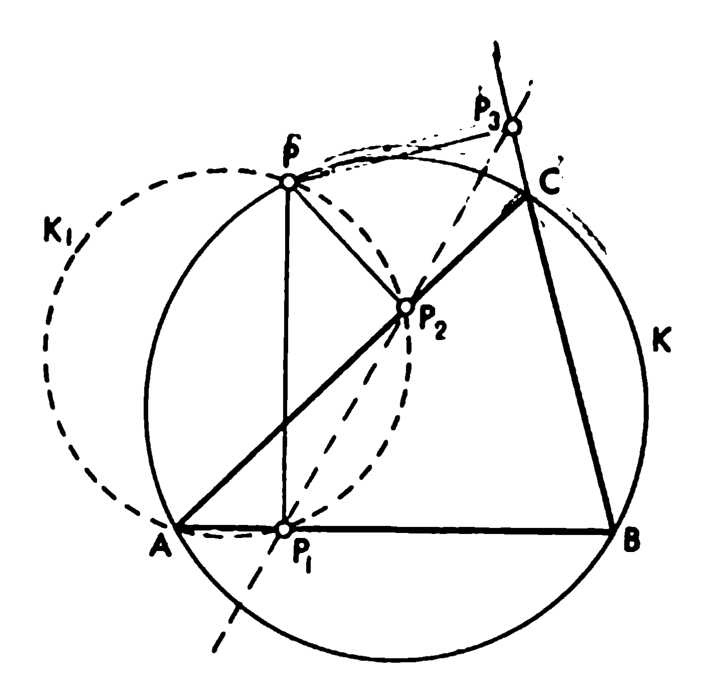


Fig. 47

18-0601

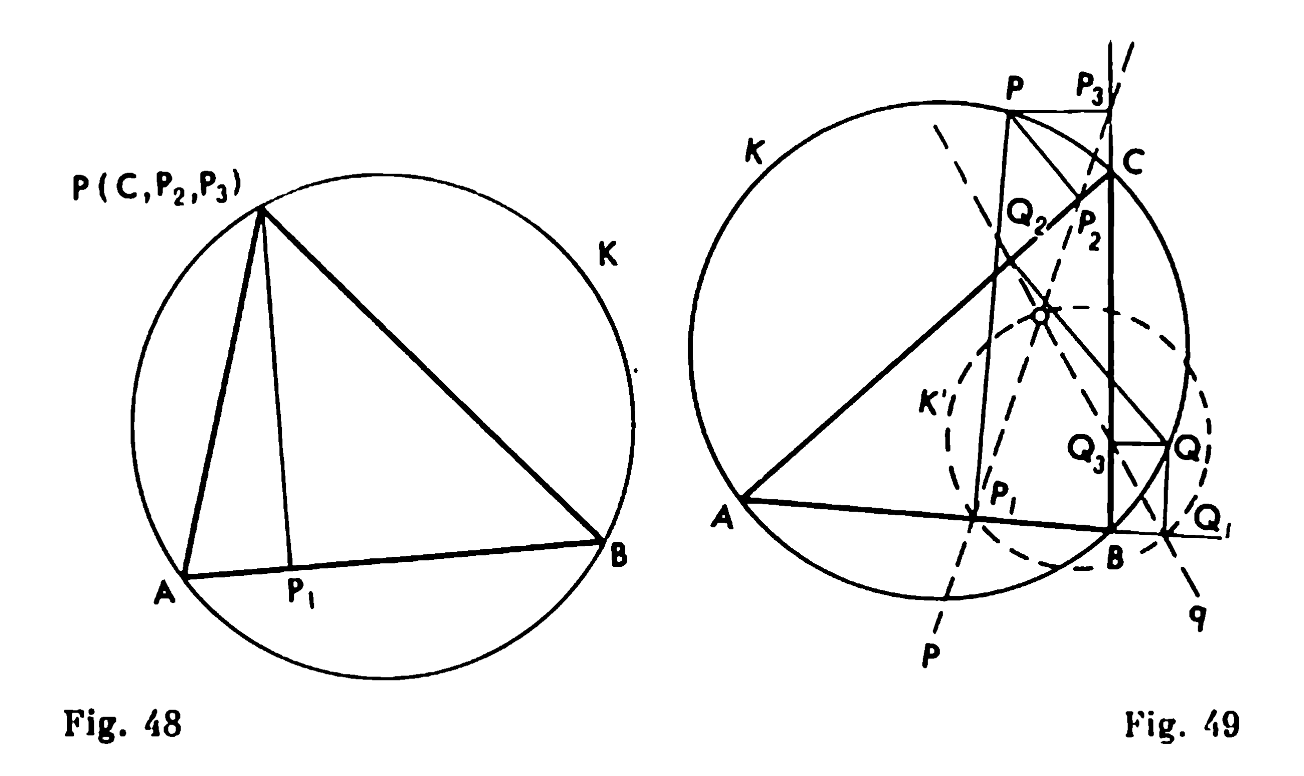


Fig. 48

Fig. 49

En vertu donc du théorème 10, leurs vitesses angulaires sont égales. Et la vitesse angulaire du rayon P_1P_2 est égale à ω. D'une façon analogue, la vitesse angulaire du rayon P_1P_3 est égale à ω.

Ainsi, les rayons P_1P_2 et P_1P_3 se déplacent à une même vitesse angulaire. Donc, l'angle qu'ils forment est constant. En vue de prouver qu'il est nul (ce qui achèvera notre démonstration) faisons coïncider le point C avec le point P (fig. 48). Le point P_2 est alors confondu avec P_3 (et avec P et C), et l'angle considéré est nul. Il est donc toujours nul.

● EXERCICE 9. Soient P et Q deux points du cercle (K) circonscrit à un triangle ABC. Montrer que le point d'intersection des droites de Simpson (p) et (q) des points P et Q (fig. 49) décrit un cercle (K') lorsque le point C décrit le cercle K (les points A, B, P et Q sont fixes).

10. Soient P un point du cercle (K) circonscrit à un triangle ABC, et P_1, P_2, P_3 les points symétriques de P par rapport aux côtés AB, AC et BC. Montrer que les points P_1, P_2 et P_3 sont situés sur une même droite qui passe par l'orthocentre du triangle ABC (fig. 50).

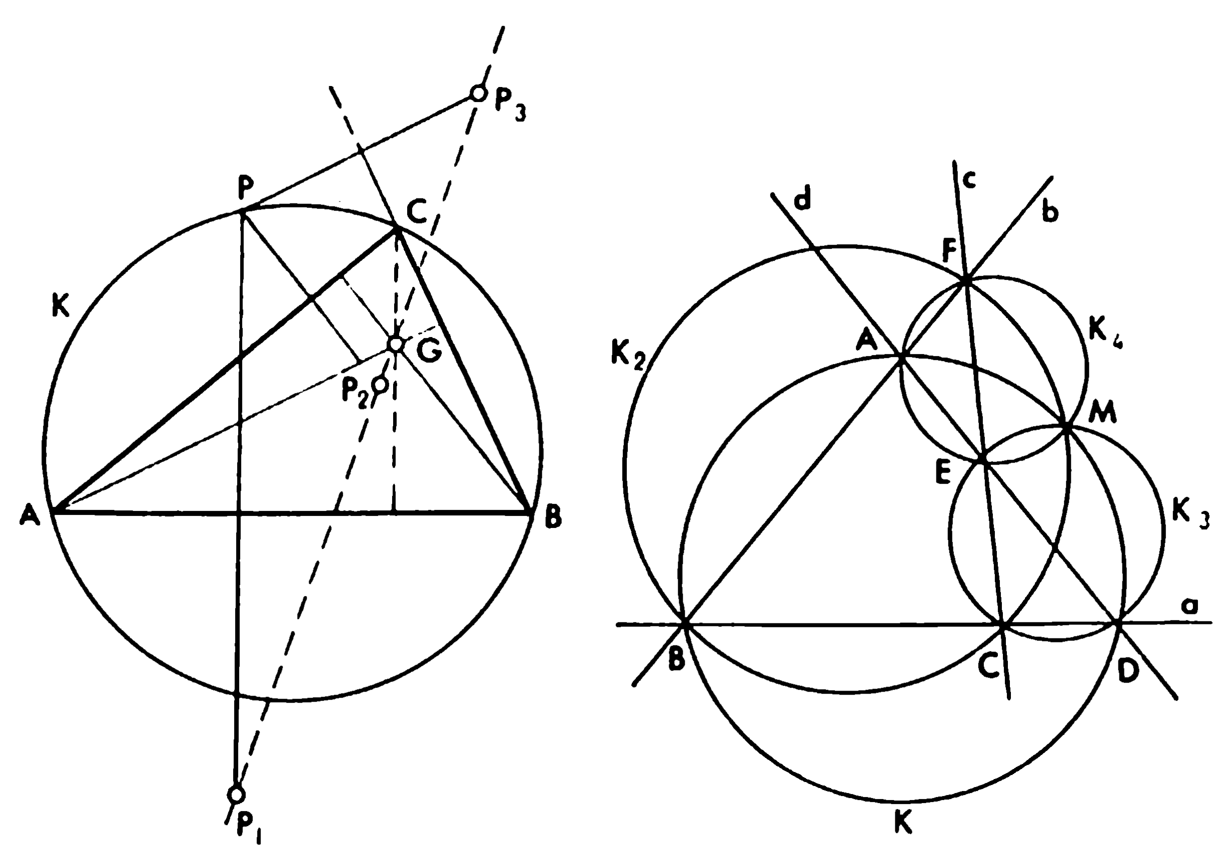


Fig. 50

● PROBLÈME 6. *Montrer que les cercles* (K_1), (K_2), (K_3) et (K_4) *circonscrits aux triangles ABD, BFC, CED et AEF, formés par l'intersection deux à deux des droites* (*a*), (*b*), (*c*) *et* (*d*) *passent par un même point* (fig. 51).

● SOLUTION. Les cercles (K_1) et (K_2) possédant un point commun B doivent en posséder un autre[1]) que nous désignerons par M. Prouvons que M appartient également aux cercles (K_3) et (K_4).

Fixons les points B, C et D et faisons pivoter les droites (*b*), (*c*) et (*d*) à la même vitesse angulaire. Comme l'angle

---

[1]) Si (K_1) et (K_2) étaient tangents en B, les triangles ABD et BFC seraient semblables, et les droites AD et FC parallèles, contrairement à l'hypothèse.

$\hat{B}A\hat{D}$ reste constant (voir théorème 10, § 2), le point de concours A des droites (b) et (d) décrira le cercle (K_1). D'une façon analogue, le point de concours F des droites (b) et (c) décrira le cercle (K_2), le point de concours E des droites (c) et (d), le cercle (K_3).

A un moment donné le point A se confondra avec le point M (voir fig. 52,a) et alors par ce point passeront les droites (b) et (d). Comme M appartient également à K_2 et que le point de concours des droites (b) et (c) est situé sur le cercle (K_2), il en résulte qu'à ce moment les droites (b), (c) et (d) se couperont en M. Or le point de concours des droites (c) et (d) appartient au cercle (K_3), donc le point M est situé sur ce cercle.

Pour démontrer que le point M appartient également au cercle (K_4), fixons les points B, A et F et faisons pivoter autour d'eux et à une même vitesse angulaire les droites (a), (d) et (c) (fig. 52,b). Comme précédemment à un moment donné les droites (a), (d) et (c) se couperont en M. Donc (K_4) passe par M.

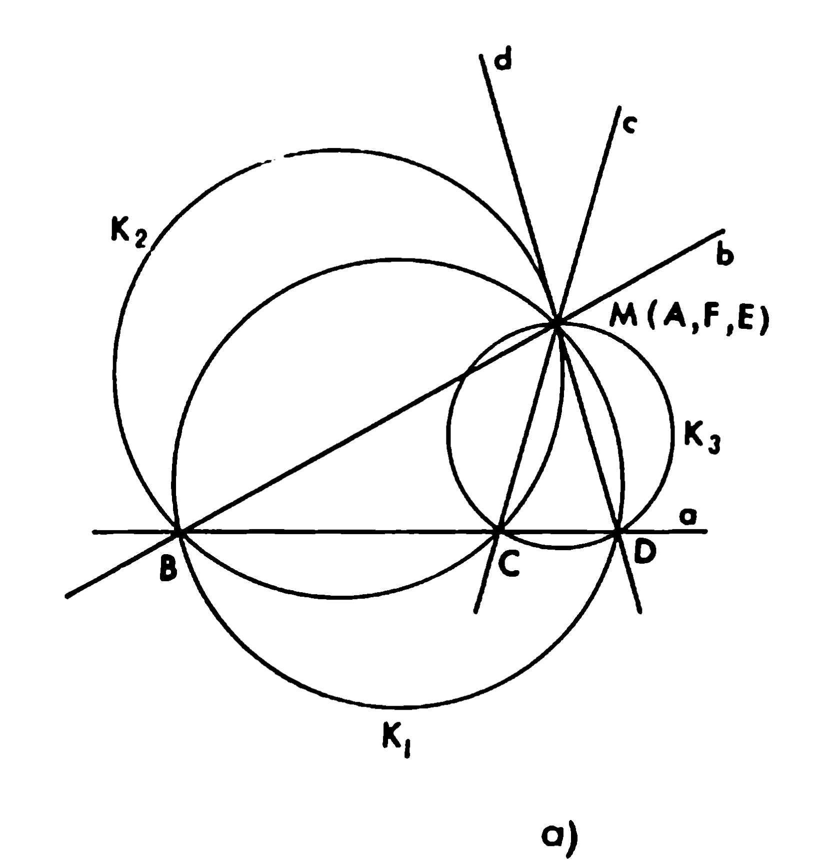
● EXERCICES.

11. Soit un point quelconque M du côté AB d'un triangle ABC. Montrer que les centres O_1, O_2 et O_3 des cercles circonscrits aux triangles ABC, AMC et BMC sont situés sur un cercle passant par C (fig. 53).

12. Montrer que les centres des cercles (K_1), (K_2), (K_3) et (K_4) du problème 6 sont situés sur un cercle qui passe par le point d'intersection de ces cercles (fig. 54).

● PROBLÈME 7. *Deux cercles* (K_1) *et* (K_2) *se coupent en A et B* (fig. 55). *Un point M décrit le cercle* (K_1). *Désignons par N et P les points d'intersection des droites MA et MB avec le cercle* (K_2). *Montrer que le lieu géométrique du centre O du cercle* (K) *circonscrit au triangle MNP est un cercle.*

● SOLUTION. Lorsque le point M décrit le cercle (K_1), les rayons AN et BP pivotent autour des points A et B à la même vitesse angulaire ω. Les vitesses angulaires des rayons O_2N et O_2P sont égales[1]) à 2 ω. Il en résulte que



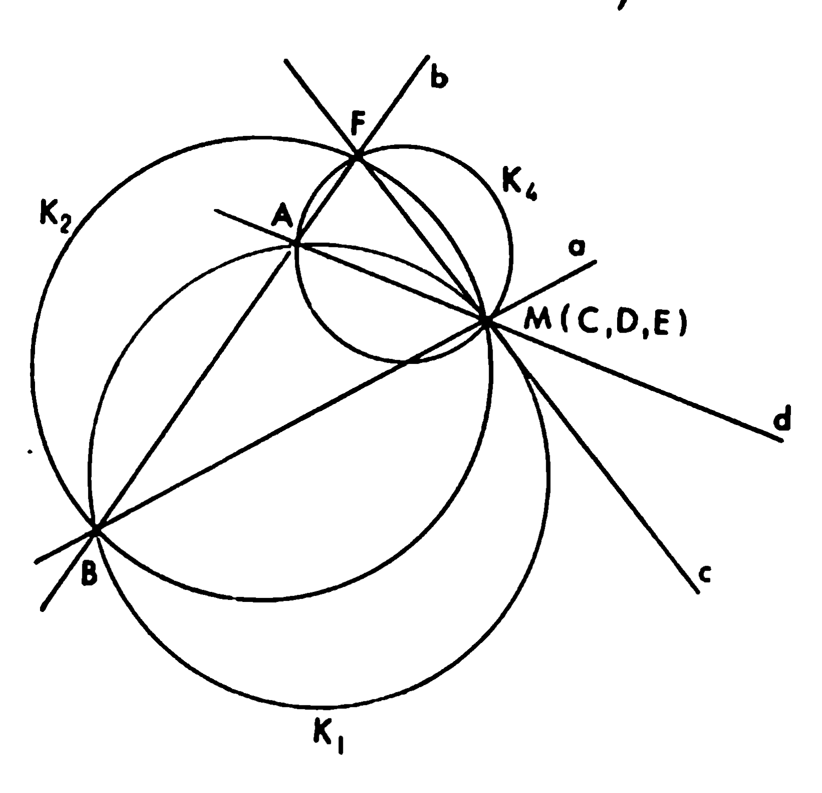


Fig. 52

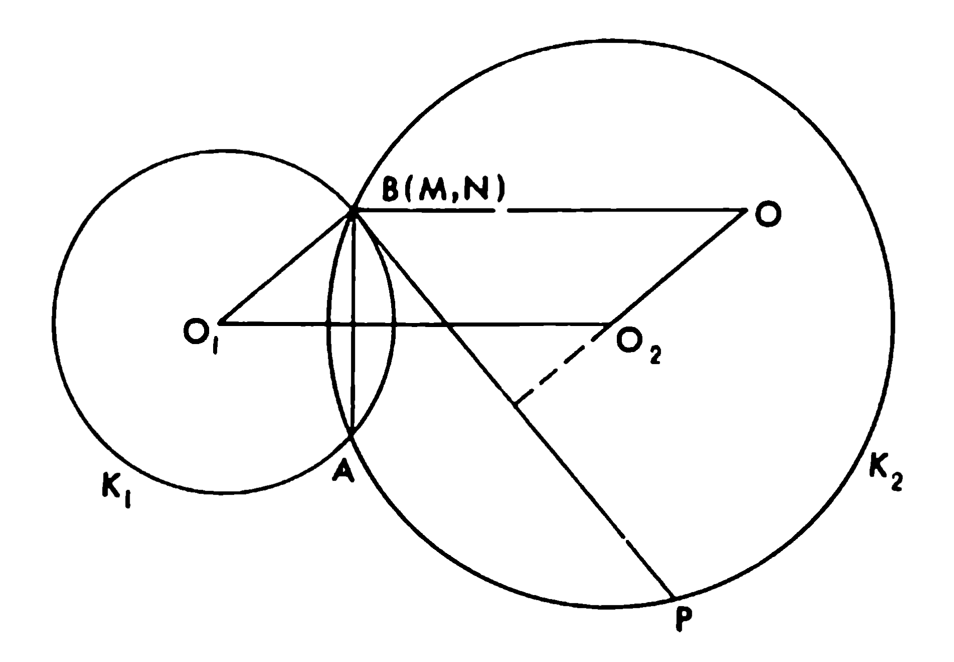


Fig. 56

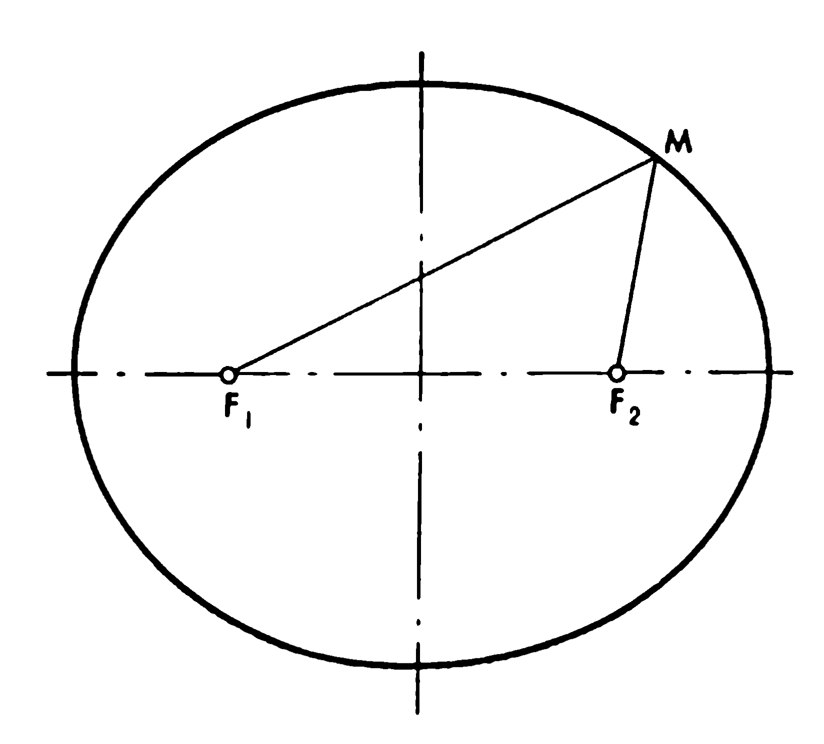


Fig. 57

tangente en B au cercle (K_1) (voir n° 2, § 2). La corde MA se confond avec la corde AB et le point N avec le point B. Le triangle MNP se réduit au segment BP. Le centre O du cercle qui lui est circonscrit est situé à l'intersection de la médiatrice de la corde BP et de la perpendiculaire en B[1]) à la corde AB. Le quadrilatère O_1O_2OB est donc un parallélogramme et $O_2O = R_1$.

Ainsi le lieu géométrique du point O est un cercle de centre O_2 et de rayon R_1.

● EXERCICES.

13. Montrer que le côté PN du triangle MNP (cf. problème 7) est tangent à un cercle fixe.

14. Montrer que le lieu géométrique de l'orthocentre du triangle MNP du problème 7 est un cercle.

En conclusion nous allons étudier quelques propriétés de l'ellipse, l'hyperbole et la parabole.

L'*ellipse* est le lieu géométrique des points dont la somme des distances a deux points fixes F_1 et F_2 (appelés *foyers*) est une grandeur constante (fig. 57).

● PROBLÈME 8. *Montrer que la tangente en un point M d'une ellipse forme les angles égaux avec les rayons vecteurs $\overrightarrow{MF_1}$ et $\overrightarrow{MF_2}$ du point M*[2]). *Et réciproquement. Etant donnés deux points fixes F_1 et F_2 et une courbe, montrer que si la tangente en chaque point M de cette courbe forme des angles égaux avec les rayons vecteurs MF_1 et MF_2, alors cette courbe est une ellipse (ou un arc d'ellipse) de foyers F_1 et F_2.*

● SOLUTION. Supposons que le point M se déplace sur l'ellipse à une vitesse **v**. La projection du vecteur **v** sur les

---

[1]) Le côté MN du triangle MNP a dégénéré en un point, mais lorsqu'on passe à la limite, son sens se confond avec celui de la corde AB.

[2]) Ce problème (qui en fait est un théorème) se formule aussi comme suit: montrer que la tangente en un point quelconque M d'une ellipse est bissectrice extérieure de l'angle des rayons vecteurs du point M.

rayons vecteurs[1]) $\mathbf{r}_1 = \overrightarrow{MF_1}$ et $\mathbf{r}_2 = \overrightarrow{MF_2}$ (fig. 58) est

$$\mathbf{v}_1 = \text{proj}_{\mathbf{r}_1} \mathbf{v} = \mathbf{v} \cos \alpha, \qquad \mathbf{v}_2 = \text{proj}_{\mathbf{r}_2} \mathbf{v} = -\mathbf{v} \cos \beta, \quad (8)$$

où α et β désignent les angles de $\mathbf{r}_1$ et $\mathbf{r}_2$ avec la tangente. Comme, par définition de l'ellipse $|\mathbf{r}_1| + |\mathbf{r}_2| = \text{const}$, il vient[2]):

$$\mathbf{v}_1 + \mathbf{v}_2 = 0$$

et, compte tenu de (8),

$$\mathbf{v} \cos \alpha - \mathbf{v} \cos \beta = 0$$

ou encore

$$\cos \alpha = \cos \beta$$

d'où $\alpha = \beta$ puisque ces angles sont aigus.

Réciproquement, supposons que la tangente en M à une courbe L forme des angles égaux avec les rayons vecteurs $\overrightarrow{MF_1}$ et $\overrightarrow{MF_2}$ (F_1 et F_2 étant deux points fixes). Projetons le vecteur vitesse $\mathbf{v}$ du point M sur les rayons vecteurs de ce point :

$$\mathbf{v}_1 = \text{proj}_{\mathbf{r}_1} \mathbf{v} = \mathbf{v} \cos \alpha, \qquad \mathbf{v}_2 = \text{proj}_{\mathbf{r}_2} \mathbf{v} = -\mathbf{v} \cos \alpha,$$

où α désigne l'angle de la tangente en M avec les rayons vecteurs de M. En ajoutant ces relations on obtient :

$$\mathbf{v}_1 + \mathbf{v}_2 = 0,$$

d'où il résulte que la somme des modules des vecteurs $\mathbf{r}_1$ et $\mathbf{r}_2$ est une constante. Donc, la courbe L est une ellipse.

L'*hyperbole* est le lieu géométrique des points dont la valeur absolue de la différence à deux points fixes F_1 et F_2, appelés foyers, est une grandeur constante (fig. 59).

---

[1]) C'est-à-dire sur les supports de ces vecteurs.

[2]) La somme $|\mathbf{r}_1| + |\mathbf{r}_2|$ étant une grandeur constante, les accroissements des vecteurs $\mathbf{r}_1$ et $\mathbf{r}_2$ sont égaux en valeur absolue et de sens contraires. Donc les vitesses de ces accroissements sont égales en valeur absolue et de sens contraires. En vertu du nº 5 § 2, elles sont égales à $\mathbf{v}_1$ et $\mathbf{v}_2$.

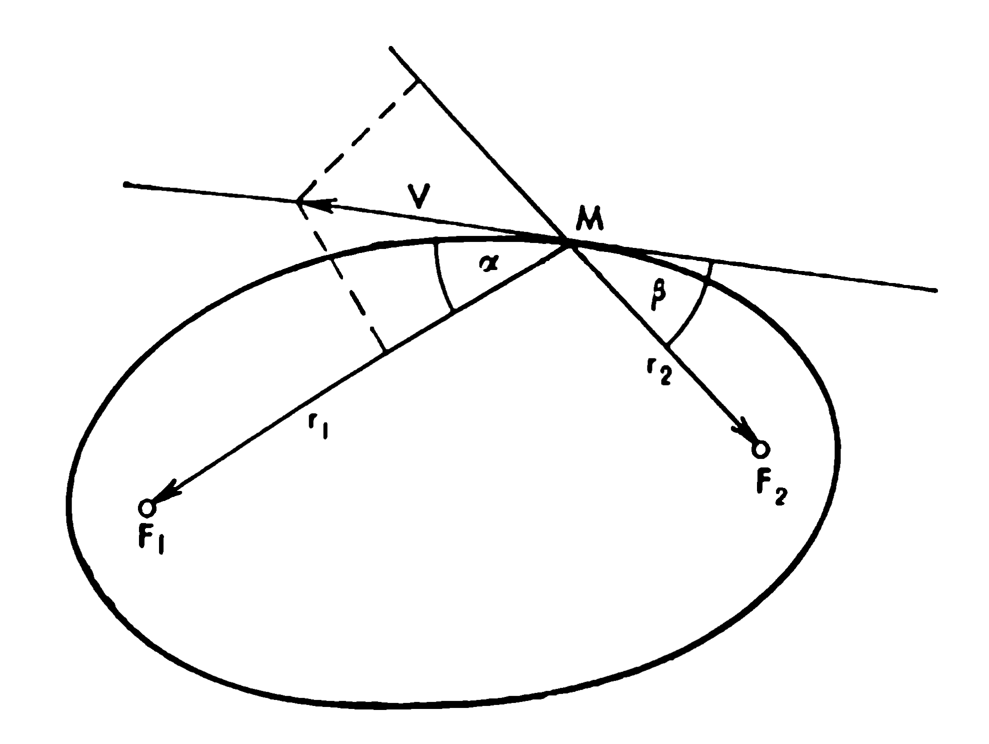


Fig. 58

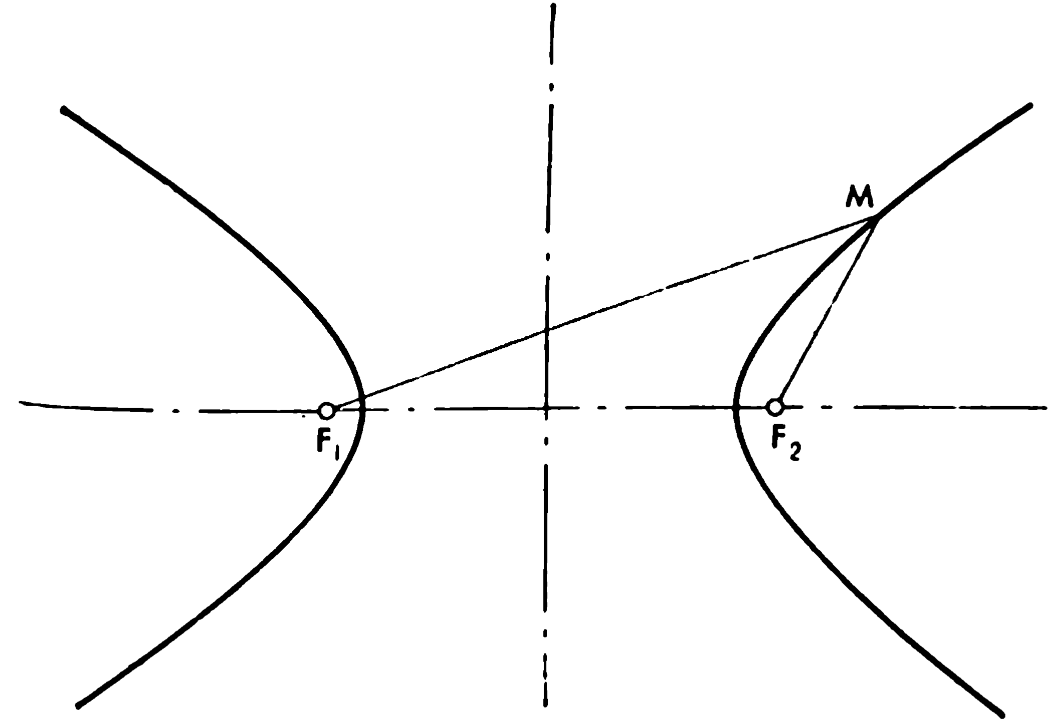


Fig. 59

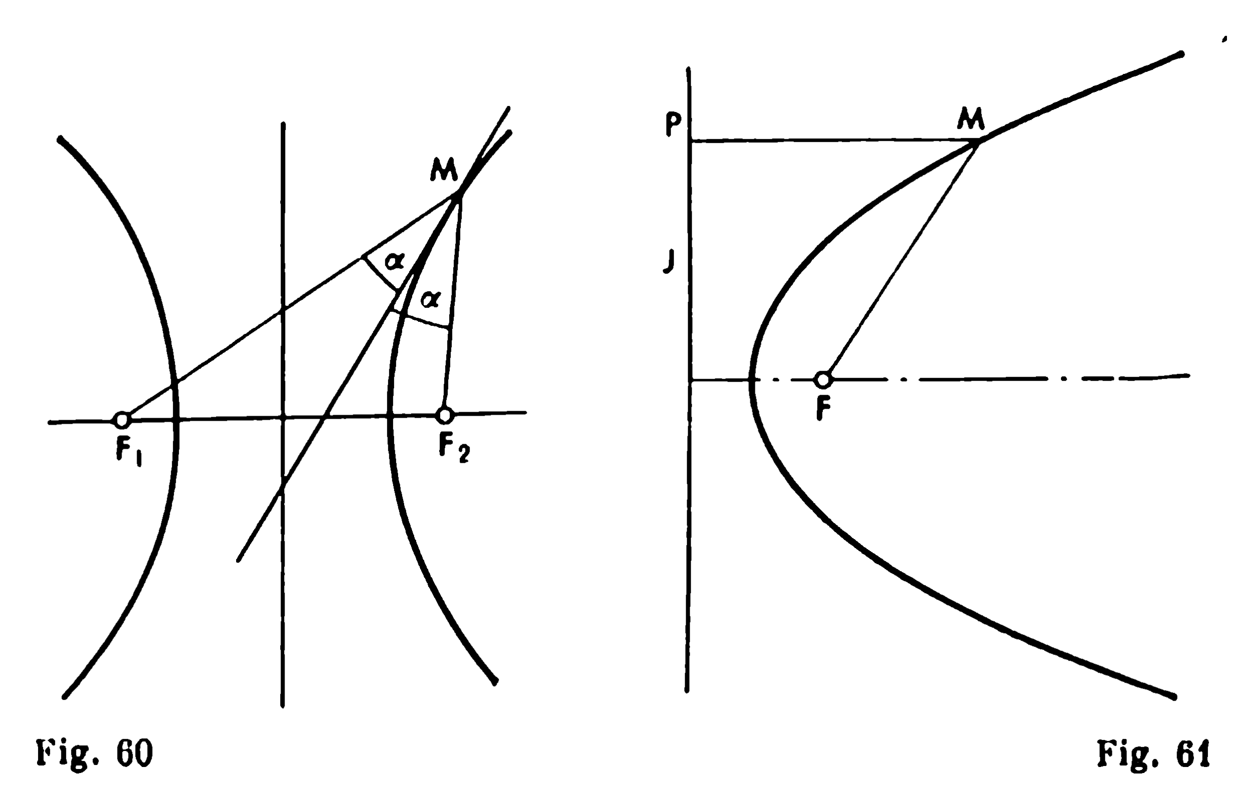


Fig. 60 Fig. 61

● EXERCICE. 15. Montrer que la tangente à une hyperbole est bissectrice intérieure de l'angle des rayons vecteurs du point de tangence (fig. 60).

Et réciproquement. Si la tangente en tout point M d'une courbe L est bissectrice intérieure de l'angle des rayons vecteurs $\overrightarrow{MF_1}$ et $\overrightarrow{MF_2}$ (F_1 et F_2 sont des points fixes) du point M, la courbe L est une hyperbole (ou une branche d'hyperbole) de foyers F_1 et F_2.

La *parabole* est le lieu géométrique des points situés à égale distance d'un point fixe F appelé foyer et d'une droite donnée (d), ne passant pas par F, appelée directrice (fig. 61).

● EXERCICE. 16. Montrer que la tangente en un point M d'une parabole est bissectrice intérieure de l'angle des rayons vecteurs [1]) du point M.

Et réciproquement. Etant donnés une droite (d) et un point fixe F situé hors d'elle, si la tangente en un point M d'une courbe est bissectrice intérieure de l'angle des rayons vecteurs $\overrightarrow{MF}$ et $\overrightarrow{MP}$ (P étant

[1]) Dans le cas de la parabole les rayons vecteurs d'un point M sont MF et MP, P désignant le pied de la perpendiculaire abaissée de M sur la directrice (d).

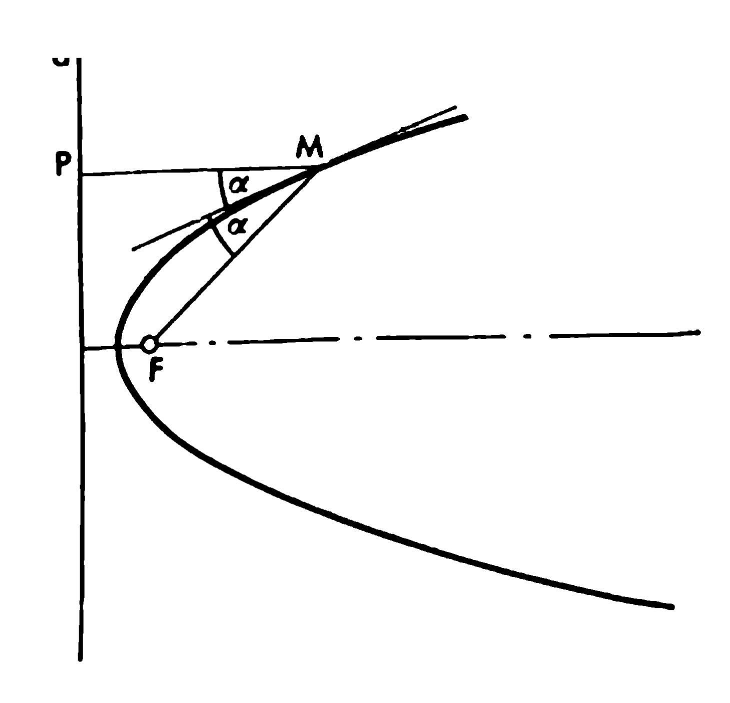


Fig. 62

le pied de la perpendiculaire abaissée de M sur (d)), alors cette courbe est une parabole (ou un arc de parabole) de foyer F et de directrice (d) (fig. 62).

Indications

1. Même solution que pour le problème 1. Il importe de remarquer que les points O_1, O_2, O_3 se confondent (le triangle $O_1O_2O_3$ se réduit donc à un point) et les vecteurs $\overrightarrow{O_1O_2}$ et $\overrightarrow{O_1O_3}$ s'annulent.

La relation

$$\overrightarrow{O_1O_3} = U_{\frac{\pi}{3}} \overrightarrow{O_1O_2} + \mathbf{R}$$

entraîne $\mathbf{R} = 0$.

2. Fixer les points A, B et faire déplacer le point C. Etablir les vitesses de points A' et B'. Dans la position déterminante, faire coïncider C et C'.

3. Même solution que pour le problème 1. Faire coïncider C avec A ou B.

4. Même solution que pour le problème 3.

5. Fixer les points A et B et déplacer C et D. Les points C et D doivent se déplacer de telle sorte que le triangle CSD, dont le sommet S est fixe, reste toujours isocèle. Montrer que $\mathbf{v}_Q = \mathbf{v}_T$.

6. Comme dans l'exercice précédent, montrer que $\mathbf{v}_Q = -\mathbf{v}_T$.

7. Comme dans le problème 1, montrer que

$$\overrightarrow{AA'} = U_{-\frac{\pi}{3}}\overrightarrow{C'C}, \quad \overrightarrow{BB'} = U_{\frac{\pi}{3}}\overrightarrow{C'C}.$$

En déduire que les points d'intersection deux à deux des droites AA', BB' et CC' sont situés sur le cercle circonscrit au triangle ABC.

8. Fixer les droites (b), (c) et (d) et animer la droite (a) d'un mouvement uniforme de translation. Montrer ensuite que

$$v_{O_3} = \lambda v_{O_1}, \quad v_{O_4} = \mu v_{O_1},$$

où λ = const, μ = const. Etablir la nullité de R_1 et R_2 en considérant deux (!) positions déterminantes: lorsque (a) passe par A et lorsqu'elle passe par B. Utiliser le fait qu'un vecteur simultanément colinéaire avec deux droites non concourantes est nul.

9. Lorsque le point C se déplace, la vitesse angulaire des droites de Simpson (p) et (q) est égale à la vitesse angulaire des rayons AC et BC.

10. En utilisant les résultats du problème 5, montrer d'abord que les points P_1, P_2, P_3 sont alignés. Puis, après avoir fixé les points A, B et P, faire pivoter à la même vitesse les droites AC et BC autour des points A et B. Etudier les vitesses angulaires des droites $P_1P_2P_3$ et P_1G. Choisir la position déterminante comme dans le problème 5.

11. Fixer les points A et C et faire pivoter autour d'eux et à la même vitesse angulaire les droites AB, CM et CB. Etudier le mouvement des points O_1, O_2 et O_3. Etudier la position de coïncidence des droites AC et AB.

12. Fixer les points B, C et D et faire pivoter autour d'eux et à la même vitesse angulaire les droites BF, CF et DA jusqu'à ce qu'elles passent par le point M. Se servir du résultat de l'exercice précédent.

13. Se servir du fait que la corde PN est invariante (cf. solution du problème 7).

14. Démontrer préalablement que la distance d'un sommet à l'orthocentre du triangle est le double de la distance du centre du cercle circonscrit au côté opposé au sommet considéré. Cette démonstration est facile à faire sans recourir à la cinématique. Se servir du fait que lorsque les points M, N et P se déplacent, les droites O_1M et NP restent perpendiculaires et la distance du point O à la droite NP reste constante.

15. Même solution que pour le problème 8.

16. Même solution que pour le problème 8, sauf qu'au lieu de considérer la vitesse de variation du module du rayon vecteur $\overrightarrow{MF}$, on considérera celle du module du rayon vecteur $\overrightarrow{MP}$.

TABLE DES MATIÈRES

